THE JOURNAL OF EDUCATIONAL RESEARCH

EXECUTIVE EDITORS

WILSON B. THIEDE, Professor of Education, Educational Administration, and Agricultural and Extension Education, The University of Wisconsin. Chairman.

THEODORE HARRIS, Professor of Educational Psychology and Education, The University of Wisconsin.

L. JOSEPH LINS, Professor and Coordinator of Institutional Studies, The University of Wisconsin.

PHILIP LAMBERT, Professor of Educational Psychology and Educational Administration, The University of Wisconsin.

JOHN SCHMID, Professor of Education and Chairman of the Department of Educational Foundations, University of Arkansas.

VOLUME LVI
September 1962 - August 1963

BUREAU OF EDNL. & PSYL. RESEARCH : : LIBRARY

Copyright 1962-1963 by

DEMBAR EDUCATIONAL RESEARCH SERVICES, Inc.
Madison 3, Wisconsin

Bure Edn. [illegible]
Dated 1.9.70
No. J301

INDEX

Listed Under Articles and Authors

Ability of Poor Readers to Discriminate Paired Associates Under Differing Conditions of Confirmation, Wayne Otto, Apr. 63, 428-31.
Academic Achievement of Academically Talented Students, Frederick R. Smith, Jan. 63, 255-9.
Academic Achievement and Adjustment of Underage and Overage Third Graders, Marian L. Carroll, Apr. 63, 415-19.
Age, IQ, and Improvement of Reading, Norman M. Chansky, Apr. 63, 439.
Aiken, Lewis, R., Jr. Personality Correlates of Attitude Toward Mathematics, May-June 63, 476-83.
Albert, Frank R., Jr. Selected Characteristics of School Board Members and Their Attitudes Toward Criticism of the Public Schools, Sept. 62, 55-6.
Amidon, Edmund and Blumberg, Arthur. Student Perceptions of Group Work in Teacher Education Classes, Mar. 63, 352-7.
Analysis of Remedial Activities Used by Elementary Teachers in Coping with Classroom Behavior Problems, Donald L. Barnes, July-Aug. 63, 544-7.
Analysis of the Quality of the Behaviors of Principals as Observed and Reported in Six Critical Incident Studies, Bernice Cooper, Apr. 63, 410-14.
Attitudes and Beliefs as Sources of Vocational Preference, A. G. Sorenson and Irma E. Morris, Sept. 62, 20-7.

Baird, Ruth Cates. A Survey of Errors in English Compositions, Jan. 63, 228-35.
Barnes, Donald L. An Analysis of Remedial Activities Used by Elementary Teachers in Coping with Classroom Behavior Problems, July-Aug. 63, 544-7.
Barwick, Janice M. and Arbuckle, Dugald S. A Study of the Relationship Between Parental Acceptance and the Academic Achievement of Adolescents, Nov. 62, 148-51.
Behavior Problems of Children as Viewed by Children in the Fifth through Eighth Grades, Thomas Mangan and David Shafer, Oct. 62, 104-6.
Bilingualism in Southwest Louisiana, Harold D. Love, Nov. 62, 144-7.
Bledsoe, Joseph C. The Relation of Mental Health Analysis Scores to Teacher Ratings of Mental Health Status of Elementary Pupils, May-June 63, 488-91.
Bloomer, Richard H. The Relation of Question Di ficulty to Question Form, Jan. 63, 250-4.
Blumberg, Arthur and Amidon, Edmund. Teache Reactions to School Faculty Meetings, May-June 63, 466-70.

Can the School Curriculum Make a Frontal Approach to Mental Health? Gail M. Inlow, Apr 63, 395-402.
Cappa, Dan and Schubert, Delwyn G. Do Parents Help Gifted Children Read? Sept. 62, 33-6.
Capps, Lelon Roger. A Comparison of the Common Denominator and Inversion Method in Teaching Division of Fractions, July-Aug. 6 516-22.
Carroll, Marian L. Academic Achievement and Ac justment of Underage and Overage Third Graders, Apr. 63, 415-19.
Case Studies in the Changing Roles and Statuses of College Teachers, James Moss, Sept. 62, 41-4.
Changes in Self-Awareness During the High School Years: A Study of Mental Health Using Pape and-Pencil Tests, John E. Nickols, Jr. Apr 63, 403-9.
Chansky, Norman M. Age, IQ, and Improvement of Reading, Apr. 63, 439.
Chansky, Norman M. Elementary School Teacher Rate Report Cards, July-Aug. 63, 523-8.
Claye, Clifton M. Lola Gets what Lola Wants Fro Supervision, Mar. 63, 358-61.
Comparative Effects of Small-Group vs. Teacher-Led Discussion Sessions upon Student Achiev ment and Perception in Educational Psycholo gy, William J. Gnagey, Sept. 62, 28-32.
Comparison of the Achievement of Multi-Graded and Single-Graded Rural Elementary School Children, Carmen J. Finley and Jack M. Thompson, May-June 63, 471-5.
Comparison of the Common Denominator and Inversion Method in Teaching Division of Fractions Lelon Roger Capps, July-Aug. 63, 516-22.
Comparison and Evaluation of Two Methods of Organization for the Teaching of Reading, Reginald Kierstead, Feb. 63, 317-21.
Comparison of Student Reaction to Similar, Concurrent Teacher-Content Instruction, Shirley Ul man Wedeen, July-Aug. 63, 540-3.
Comparison of the Performance of Pre-Graded Students with Grade-Motivated Students, T. L. Hawk and L. M. DeRidder, July-Aug. 63, 548-

50.
Comparisons Between Second-, Third-, and Fourth-Grade Children in the Audio-Lingual Learning of Japanese as a Second Language, Robert E. Grinder et al., Dec. 62, 191-7.
Comparison of Two Methods for Determining Difficulty in a Multiple Choice Test, Margery V. Hawes and Gerald C. Helmstadter, Apr. 63, 420-3.
Comprehension Subtest of the Wechsler Adult Intelligence Scale as an Indicator of Social Awareness, Robert M. Roth, Mar. 63, 387-90.
Cooper, Bernice. An Analysis of the Quality of the Behaviors of Principals as Observed and Reported in Six Critical Incident Studies, Apr. 63, 410-14.
Corotto, Loren V. The Prediction of Success in Initial College Mathematics, Jan. 63, 268-71.
Creaser, James W. Evaluation of a College Habits Course Using Scores on a Q-Sort Test as the Criterion, Jan. 63, 272-4.

Drumheller, Sidney J. The Image of a Teachers College as Seen by High-School Counselors, Jan. 63, 260-3.
Degree of Teacher Domination in Group Processes and Student Attitude Toward Instruction in Secondary Methods Classes, Kenneth H. Hoover, Mar. 63, 379-81.
Development of a Parent Attitude Toward Education Scale, Gene R. Medinnus, Oct. 62, 100-3.
Development of and Reasons for Vocational Choices of Adolescents Through the High-School Years, Marvin Powell and Viola Bloom, Nov. 62, 126-33.
Development of Responsibility Perceptions of Lower and Middle Class Children, Michael Zunich, May-June 63, 497-9.
Dickinson, Donald J. and Larson, J. Donald. The Effects of Chronological Age in Months on School Achievement, May-June 63, 492-3.
Differences in Achievement Test Scores Associated with Item Exposure Time and Audio Reinforcement, R. P. Kropp et al., Dec. 62, 214-17.
Discovery of Personality Traits, John L. Horn, May-June 63, 460-5.
Do Parents Help Gifted Children Read? Dan Cappa and Delwyn G. Schubert, Sept. 62, 33-6.
Docter, Robert L. The Effectiveness of Workbooks in the Elementary School Reading Program, Feb. 63, 327-9.

Educators' Attitudes Toward Educational Practices, Robert S. Miller, Apr. 63, 424-7.
Effect of Musical Distraction upon the Performance Efficiency of Children, Alvin Engel and Cynthia Engel, Sept. 62, 45-7.
Effect on Typewriting Performance of Varying the Degree of Interest in Timed-Writing Copy Materials, Willard H. Nelson, et al., Dec. 62, 221-2.
Effect of Visual Imagery Upon Spelling Performance, Leon D. Radaker, Mar. 63, 370-2.
Effects of Chronological Age in Months on School Achievement, Donald J. Dickinson and J. Donald Larson, May-June 63, 492-3.
Effects of Self-Directed Study on Course Achievement, Retention and Curiosity, Donald E. Hovey et al., Mar. 63, 346-51.
Effectiveness of Workbooks in the Elementary School Reading Program, Robert L. Docter, Feb. 63, 327-9.
Effectiveness of Different Methods of Study, Mabel S. Noall, Sept. 62, 51-2.
Elementary School Teachers Rate Report Cards, Norman M. Chansky, July-Aug. 63, 523-8.
Engel, Alvin and Engel, Cynthia. The Effect of Musical Distraction Upon the Performance Efficiency of Children, Sept. 62, 45-7.
Evaluation Based Upon Observation of Reports of Supervisors of Student Teaching, Rachel D. Wilkinson, Jan. 63, 264-7.
Evaluation of the California Achievement Test, Elementary Form W, Reading Vocabulary, Charles Sheppard and William J. Campbell, May-June 63, 481-3.
Evaluation of a College Study Habits Course Using Scores on a Q-Sort Test as the Criterion, James W. Creaser, Jan. 63, 272-4.
Evaluation of the Vocational Counseling Program of the Employment Security Service in Kanawha County, William V. Wagner, Mar, 63, 373-5.
Evaluating Teaching Personality Before Student Teaching Begins, Lebern N. Miller, Mar. 63, 382-4.
Experiment in Teaching Elementary School Methods via Closed Circuit Television, Herbert P. Rumford, Nov. 62, 139-43.

Factorial Analysis of Fields of Scholarly Production, Robert H. Knapp and Linda A. Ehlinger, May-June 63, 484-7.
Factors Affecting the Academic Underachievement of Bright High-School Students, Helen Erskine Roberts, Dec. 62, 175-83.
Factors Associated with the Progress of Students Enrolled in College Reading Program, J. Wesley Schneyer, Mar. 63, 340-5.
Factors Involved in the Adjustment of Children Entering First Grade, Sarah Lou Hammond and Dora S. Skipper, Oct. 62, 89-95.
Finley, Carmen J. and Thompson, Jack M. A Comparison of the Achievement of Multi-Graded and Single-Graded Rural Elementary School Children, May-June 63, 471-5.
Formula for Predicting the Comprehension Level of Material to be Presented Orally, John R. Rogers, Dec. 62, 218-20.
Fowler, H. Seymour. Grade Placement of High-School Biology, Nov. 62, 156-9.

Glidewell, John C. et al. Screening in Schools for Behavior Disorders: Use of Mothers' Reports of Symptoms, July-Aug. 63, 508-15.
Gnagey, William J. The Comparative Effects of Small-Group vs. Teacher-Led Discussion Sessions upon Student Achievement and Perception in Educational Psychology, Sept. 62, 28-32.
Gordon, Ira J. and Wood, Patricia C. The Relationship Between Pupil Self-Evaluation, Teacher Evaluation of the Pupil and Scholastic Achievement, Apr. 63, 441-3.
Grade Placement of High-School Biology, H. Seymour Fowler, Nov. 62, 156-9.
Grinder, Robert E. et al. Comparisons Between Second-, Third-, and Fourth-Grade Children in the Audio-Lingual Learning of Japanese as a Second Language, Dec. 62, 191-7.

Haas, Kurt. Personality Needs of Academically Superior Students and Their Parents, Mar. 63, 389-90.
Hammond, Sarah Lou and Skipper, Dora S. Factors Involved in the Adjustment of Children Entering First Grade, Oct. 62, 89-95.
Harris, Theodore L. Some Issues in Beginning Reading Instruction, Sept. 62, 5-19.
Harris, Theodore L. Summary of Investigations Relating to Reading, July 1, 1961 to June 30, 1962, Feb. 63, 283-310.
Hawes, Margery V. and Helmstadter, Gerald C. A Comparison of Two Methods for Determining Difficulty in a Multiple Choice Test, Apr. 63, 420-3.
Hawk, T. L. and DeRidder, L. M. A Comparison of the Performance of Pre-Graded Students with Grade-Motivated Students, July-Aug. 63, 548-50.
Herman, Wayne L. How Intermediate Children Rank the Subjects, Apr. 63, 435-6.
Hildreth, Gertrude. Simplified Handwriting for Today, Feb. 63, 330-3.
Hoover, Kenneth H. Degree of Teacher Domination in Group Processes and Student Attitude Toward Instruction in Secondary Methods Classes, Mar. 63, 379-81.
Horn, John L. The Discovery of Personality Traits, May-June 63, 460-5.
Hovey, Donald et al., Effects of Self-Directed Study on Course Achievement, Retention and Curiosity, Mar. 63, 346-51.
How Intermediate Children Rank the Subjects, Wayne L. Herman, Jr., Apr. 63, 435-8.
How Much Are Students in the Public Schools Getting from Their School Work?, Doyne M. Smith, Nov. 62, 117-25.

Inlow, Gail. Can the School Curriculum Make a Frontal Approach to Mental Health? Apr. 63, 395-402.
Image of a Teachers College as Seen by High-School Counselors, Sidney J. Drumheller, Jan. 63, 260-3.
Intelligence and Gain in Reading as Related to Gains in the Sub-tests of the Stanford Achievement Test, Carrie M. Scott, May-June 63, 494-6.
Investigation of the Attitudes of Parents of Unsuccessful and Successful Readers, Dorothy P. Macdonald, Apr. 63, 437-8.
Investigation of the Differences Between Tenth-, Eleventh-, and Twelfth-Grade "Conforming" and "Nonconforming" Boys, Marvin Powell and Jerry Bergem, Dec. 62, 184-90.
Investigation of the Relationship of Some Pre-Adolescent Developmental Factors to Adolescent Social Adjustment, Israel Woronoff, Nov. 62, 164-6.

Kamii, Constance K. and Weikart, David P. Marks, Achievement, and Intellgence of Seventh Graders Who Were Retained (Nonpromoted) Once in Elementary School, May-June 63, 452-9.
Kierstead, Reginald. A Comparison and Evaluation of Two Methods of Organization for the Teaching of Reading, Feb. 63, 317-21.
Kingston, Albert J. The Relationship of First-Grade Readiness to Third- and Fourth-Grade Achievement, Oct. 63, 61-7.
Knapp, Robert H. and Ehlinger, Linda A. A Factorial Analysis of Fields of Scholarly Production, May-June 63, 484-7.
Knapp, Robert H. and Ehlinger, Linda A. Styles of Scholarly Production, Jan. 63, 243-9.
Kropp, R. P. et al. Differences in Achievement Test Scores Associated with Item Exposure Time and Audio Reinforcement, Dec. 62, 214-17.
Kropp, Russell P. and Anderson, John E. Teacher Roles Before and After Internship, Mar. 63, 366-9.

Laboratory School: An Historical Perspective, Pose M. Lamb, Oct. 62, 107-9.
Lamb, Pose M. The Laboratory School: An Historical Perspective, Oct. 62, 107-9.
Lambert, Philip. The "Successful" Child: Some Implications of Teacher Stereotyping, July-August 63, 551-3.
Leton, Donald A. Personality Ratings of High-School Students, Nov. 62, 160-3.
Likes and Dislikes of Indonesian Pupils, R. Murray Thomas and Winarno Surachmad, Oct. 62, 68-74.
Lola Gets What Lola Wants from Supervision, Clifton M. Claye, Mar. 63, 358-61.
Longitudinal Study of Auditory Discrimination, Bertha Boya Thompson, Mar. 63, 376-8.
Love, Harold D. Bilingualism in Southwest Louisiana, Nov. 62, 144-7.

Macdonald, Dorothy P. An Investigation of the At-

titudes of Parents of Unsuccessful and Successful Readers, Apr. 63, 437-8.
Mangan, Thomas and Shafer, David. Behavior Problems of Children as Viewed by Children in the Fifth through Eighth Grades, Oct. 62, 104-6.
Mastin, Victor E. Teacher Enthusiasm, Mar. 63, 385-6.
Marks, Achievement and Intelligence of Seventh Graders who were Retained (Nonpromoted) Once in Elementary School, Constance K. Kamii and David P. Weikart, May-June 63, 452-9.
Marks, Edmond, et al. Scholastic Aptitudes, Vocational Interests, and Personality Characteristics of Journalism Students, Sept. 62, 37-40.
Medinnus, Gene R. The Development of a Parent Attitude Toward Education Scale, Oct. 62, 100-3.
Method for Identifying Prospective Counselees in College, Robert M. Roth, Jan. 63, 275-6.
Meyer, Margaret R. Parent-Teacher Relationships in the Early Nineteenth Century, Sept. 62, 48-50.
Miller, Lebern N. Evaluating Teaching Personality Before Student Teaching Begins, Mar. 63, 382-4.
Miller, Robert S. Educator's Attitudes Toward Educational Practices, Apr. 63, 424-7.
Miller, Wilbur R. Readability Versus Reading Ability, Dec. 62, 205-9.
Mitchell, James V. Self-Family Perceptions Related to Self-Acceptance, Manifest Anxiety, and Neuroticism, Jan. 63, 236-42.
Moss, James. Case Studies in the Changing Roles and Statuses of College Teachers, Sept. 62, 41-4.

Nelson, Willard et al. The Effect on Typewriting Performance of Varying the Degree of Interest in Timed-Writing Copy Materials, Dec. 62, 221-2.
Nickols, John E. Changes in Self-Awareness During the High School Years, A Study of Mental Health Using Paper-and-Pencil Tests, Apr. 63, 403-9.
Noall, Mabel S. Effectiveness of Different Methods of Study, Sept. 62, 51-2.

Otto, Wayne. Ability of Poor Readers to Discriminate Paired Associates Under Differing Conditions of Confirmation, Apr. 63, 428-31.
Outcomes of Curriculum Modifications Designed to Foster Critical Thinking, Norman E. Wallen, et al., July-Aug. 63, 529-34.

Parent-Teacher Relationships in the Early Nineteenth Century, Margaret R. Meyer, Sept. 62, 48-50.
Prediction of Success in Initial College Mathematics, Loren V. Corotto, Jan. 63, 268-71.
Personality Correlates of Attitude Toward Mathematics, Lewis R. Aiken, Jr., May-June 63, 476-83.
Personality Needs of Academically Superior Students and Their Parents, Kurt Haas, Mar. 63, 389-90.
Personality Ratings of High-School Students, Donald A. Leton, Nov. 62, 160-3.
Powell, Marvin and Bergem, Jerry. An Investigation of the Differences Between Tenth-, Eleventh-, and Twelfth-Grade "Conforming" and "Nonconforming" Boys, Dec. 62, 184-7.
Powell, Marvin and Bloom, Viola. Development of and Reasons for Vocational Choices of Adolescents Through the High-School Years, Nov. 62, 126-33.

Radaker, Leon D. The Effect of Visual Imagery Upon Spelling Performance, Mar. 63, 370-2.
Readers Who Lacked Word Analysis Skills: A Group Diagnosis, Mildred C. Robeck, Apr. 63, 432-4.
Readability Versus Reading Ability, Wilbur R. Miller, Dec. 62, 205-9.
Relation of Mental Health Analysis Scores to Teacher Ratings of Mental Health Status of Elementary Pupils, Joseph C. Bledsoe, May-June 63, 488-91.
Relation Between Junior High-School Students' Problems and Parental Attitude Toward Child Rearing and Family Life, Michael Zunich, Nov. 62, 134-8.
Relation of Question Difficulty to Question Form, Richard H. Bloomer, Jan. 63, 250-4.
Relationship Between Intelligence Quotients and Gain in Reading Achievement with Arithmetic Reasoning, Social Studies and Science, Carrie M. Scott, Feb. 63, 322-6.
Relationship Between Pupil Self-Evaluation, Teacher Evaluation of the Pupil and Scholastic Achievement, Ira J. Gordon and Patricia C. Wood, Apr. 63, 441-3.
Relationship of First-Grade Readiness to Third- and Fourth-Grade Achievement, Albert J. Kingston, Oct. 62, 61-7.
Relationship of Selected Adjustment Factors, College Ability, and Achievement of Drop-Outs and Nondrop-Outs of College Freshmen, Warren J. Vorreyer, Mar. 63, 362-5.
Robeck, Mildred C. Readers Who Lacked Word Analysis Skills: A Group Diagnosis, Apr. 63, 432-4.
Roberts, Helen Erskine. Factors Affecting the Academic Underachievement of Bright High-School Students, Dec. 62, 175-83.
Robinson, Alvin E. Sixth-Grade Children Evaluate Twelve Cases of Malicious Mischief, July-Aug. 63, 554-5.
Rogers, John R. A Formula for Predicting the Comprehension Level of Material to be Presented Orally, Dec. 62, 218-20.

Roth, Robert M. The Comprehension Subtest of the Wechsler Adult Intelligence Scale as an Indicator of Social Awareness, Mar. 63, 387-8.

Roth, Robert M. A Method for Identifying Prospective Counselees in College, Jan. 63, 275-8.

Rumford, Herbert P. An Experiment in Teaching Elementary School Methods via Closed Circuit Television, Nov. 62, 139-43.

Sanders, Richard M. The Use of Intelligence Tests, May-June 63, 500.

Schneyer, J. Wesley. Factors Associated with the Progress of Students Enrolled in College Reading Program, Mar. 63, 340-5.

Scholastic Aptitudes, Vocational Interests, and Personality Characteristics of Journalism Students, Edmond Marks et al., Sept. 62, 37-40.

Scott, Carrie M. Intelligence and Gain in Reading as Related to Gains in the Sub-Tests of the Stanford Achievement Test, May-June 63, 494-6.

Scott, Carrie M. The Relationship Between Intelligence Quotients and Gain in Reading Achievement with Arithmetic Reasoning, Social Studies and Science, Feb. 63, 322-6.

Screening in Schools for Behavior Disorders: Use of Mothers' Reports of Symptoms, John C. Glidewell et al., July-Aug. 63, 508-15.

Selected Characteristics of School Board Members and Their Attitudes Toward Criticisms of the Public Schools, Frank R. Albert, Sept. 62, 55-6.

Self-Family Perceptions Related to Self-Acceptance, Manifest Anxiety, and Neuroticism, James V. Mitchell, Jan. 63, 236-42.

Sheppard, Charles and Campbell, William J. An Evaluation of the California Achievement Test, Elementary Form W, Reading Vocabulary, May-June 63, 481-3.

Simplified Handwriting for Today, Gertrude Hildreth, Feb. 63, 330-3.

Sixth-Grade Children Evaluate Twelve Cases of Malicious Mischief, Alvin E. Robinson, July-Aug. 63, 554-5.

Smith, Doyne M. How Much Are Students in the Public Schools Getting from Their School Work? Nov. 62, 117-25.

Smith, Frederick R. The Academic Achievement of Academically Talented Students, Jan. 63, 255-9.

Some Illustrated Comments on Hypothesis Construction and Research, Donald J. Willower, Dec. 62, 210-13.

Some Issues in Beginning Reading Instruction, Theodore L. Harris, Sept. 62, 5-19.

Sorenson, A. G. and Morris, Irma E. Attitudes and Beliefs as Sources of Vocational Preference, Sept. 62, 20-7.

Stafford, Curt. Teacher Time Utilization with Teacher Aides, Oct. 62, 82-8.

Student Perceptions of Group Work in Teacher Education Classes, Edmund Amidon and Arthur Blumberg, Mar. 63, 352-7.

Study of the Effectiveness of Certain Techniques of Reading Instruction on Developing the Ability of Junior High-School Students to Conceptualize Social Studies Content, Mary Witt, Dec. 62, 198-204.

Study of Experimental Methods for Teaching Emotionally Disturbed, Brain Damaged, Retarded Readers, Max Talmadge et al., Feb. 63, 311-16.

Study of Morale in Education Utilizing Incomplete Sentences, John H. Suehr, Oct. 62, 75-81.

Study of the Relationship Between Parental Acceptance and the Academic Achievement of Adolescents, Janice M. Barwick and Dugald S. Arbuckle, Nov. 62, 148-51.

Styles of Scholarly Production, Robert H. Knapp and Linda A. Ehlinger, Jan. 63, 243-9.

"Successful" Child: Some Implications of Teacher Stereotyping, Philip Lambert, July-Aug. 63, 551-3.

Suehr, John H. A Study of Morale in Education Utilizing Incomplete Sentences, Oct. 62, 75-81.

Summary of Investigations Relating to Reading, July 1, 1961 to June 30, 1962, Theodore L. Harris, Feb. 63, 283-310.

Survey of Errors in English Compositions, Ruth Cates Baird, Jan. 63, 228-35.

Swartz, Phyllis et al. Value Variations in Fourth- and Seventh-Grade Students, Oct. 62, 96-9.

Talmadge, Max et al. A Study of Experimental Methods for Teaching Emotionally Disturbed, Brain Damaged, Retarded Readers, Feb. 63, 311-16.

Teacher Enthusiasm, Victor E. Mastin, Mar. 63, 385-6.

Teacher Reactions to School Faculty Meetings, Arthur Blumberg and Edmund Amidon, May-June 63, 466-70.

Teacher Roles Before and After Internship, Russell P. Kropp and John E. Anderson, Mar. 63, 366-9.

Teacher Time Utilization with Teacher Aides, Curt Stafford, Oct. 62, 82-8.

Teachers' Ratings of High-School Students on "Likability" and Their Relation to Measures of Ability and Achievement, J. R. Williams and Walter W. Knecht, Nov. 62, 152-5.

Thomas, R. Murray and Surachmad, Winarno. Likes and Dislikes of Indonesian Pupils, Oct. 62, 68-74.

Thompson, Bertha Boya. A Longitudinal Study of Auditory Discrimination, Mar. 63, 376-8.

Typing Performance as Related to Mental Abilities and Interests: A Preliminary Study, Hortense G. White, July-Aug. 63, 535-9.

Use of Intelligence Tests, Richard M. Sanders, May-June 63, 500.

Validity of Early Entrance into Kindergarten, Rosalee G. Weiss, Sept. 62, 53-4.

Value Variations in Fourth- and Seventh-Grade Students, Phyllis Swartz et al., Oct. 62, 96-9.

Vorreyer, Warren J. Relationship of Selected Adjustment Factors, College Ability, and Achievement of Drop-Outs and Nondrop-Outs of College Freshmen, Mar. 63, 362-5.

Wagner, William V. An Evaluation of the Vocational Counseling Program of the Employment Security Service in Kanawha County, Mar. 63, 373-5.

Wallen, Norman E. et al. The Outcomes of Curriculum Modifications Designed to Foster Critical Thinking, July-Aug. 63, 529-34.

Wedeen, Shirley Ullman. Comparison of Student Reaction to Similar, Concurrent Teacher-Content Instruction, July-Aug. 63, 540-3.

Weiss, Rosalee G. The Validity of Early Entrance into Kindergarten, Sept. 62, 53-4.

Wilkinson, Rachel D. Evaluation Based Upon Observation of Reports of Supervisors of Student Teaching, Jan. 63, 264-7.

Williams, J. R. and Knecht, Walter W. Teachers' Ratings of High-School Students on "Likability" and Their Relation to Measures of Ability and Achievement, Nov. 62, 152-5.

Willower, Donald J. Some Illustrated Comments on Hypothesis Construction and Research, Dec. 62, 210-13.

Witt, Mary, A Study of the Effectiveness of Certain Techniques of Reading Instruction in Developing the Ability of Junior High-School Students to Conceptualize Social Studies Content, Dec. 62, 198-204.

White, Hortense G. Typing Performance as Related to Mental Abilities and Interests: A Preliminary Study, July-Aug. 63, 535-9.

Woronoff, Israel. An Investigation of the Relationship of Some Pre-Adolescent Development Factors to Adolescent Social Adjustment, Nov. 62, 164-6.

Editorial Board

JACOB O. BACH, Head, Educational Research Bureau, Southern Illinois University, Carbondale, Illinois. Critique Editor.
WILLIAM A. BROWNELL, Dean, School of Education, University of California, Berkeley, California. Education Psychology, Learning, and Teaching.
ROBERT A. DAVIS, Professor of Educational Research, George Peabody College for Teachers, Nashville, Tennessee. Evaluative Techniques.
GUY D. McGRATH, Dean, College of Education, Arizona State University, Tempe, Arizona. Secondary Education.
HAROLD E. MITZEL, Director, Office of Research and Evaluation, Division of Teacher Education, College of the City of New York, New York City. Teacher Personnel.
WILLARD C. OLSON, Dean of College Education, University of Michigan, Ann Arbor, Michigan. Child Development.
JOHN SCHMID, Professor, College of Education, University of Arkansas, Fayetteville, Arkansas. Field News Editor.
WILSON B. THIEDE, Professor of Education, University of Wisconsin, Madison, Wisconsin. Book Review Editor.
ARTHUR E. TRAXLER, Executive Director, Educational Records Bureau, New York 32, New York. Educational Measurements.

Associate Editors

HAROLD M. ANDERSON, Associate Professor of Education, University of Colorado, Boulder, Colorado.
VERNON E. ANDERSON, Dean, College of Education, University of Maryland, College Park, Maryland.
RUDYARD K. BENT, Professor of Education, University of Arkansas, Fayetteville, Arkansas.
EMMETT ALBERT BETTS, Research Professor, University of Miami, Coral Gables, Florida
LAWRENCE P. BLUM, Assistant Director, Student Personnel Services, University of Wisconsin–Milwaukee, Milwaukee 3, Wisconsin.
J. MURRAY LEE, Chairman, Department of Elementary Education, Southern Illinois University, Carbondale, Ill.
D. WELTY LEFEVER, Professor of Education, University of Southern California, Los Angeles, California.
L. JOSEPH LINS, Professor and Coordinator, Institutional Studies, University of Wisconsin, Madison, Wisconsin.
S. E. TORSTEN LUND, Associate Professor of Education, University of California, Berkeley, California.
WILLIAM A. McCALL, Lecturer in Education, University of Miami, Coral Gables, Florida.
DELYTE W. MORRIS, President, Southern Illinois University, Carbondale, Illinois.
VALWORTH PLUMB, Chairman, Division of Education and Psychology, University of Minnesota (Duluth Branch), Duluth, Minnesota.
DAVID G. RYANS, Rand Corporation, Santa Monica, California.
LESTER B. SANDS, Professor of Education, University of California, Santa Barbara, Goleta, California.
MANDEL SHERMAN, President, and Consultant, Emotional Health Association, Los Angeles, California.
MILDRED B. STANTON, Consultant, Special Education, State Department of Education, Hartford 15, Conn.
W. W. THEISEN, Professor of Education, Marquette University, Milwaukee, Wisconsin.
M. R. TRABUE, Professor of Higher Education, University of Kentucky, Lexington, Kentucky.
J. WAYNE WRIGHTSTONE, Director, Bureau of Educational Research, Board of Education of City of New York, Brooklyn, New York.

Editors Emeriti

E. J. ASHBAUGH, Dean Emeritus, School of Education, Miami University, Oxford MR50, Ohio.
HARRY J. BAKER, 19050 Wiltshire, Lathrup Village, Michigan.
FOWLER D. BROOKS, Director Emeritus, Department of Psychology, DePauw University, Greencastle, Ind.
LEO J. BRUECKNER, Professor Emeritus, University of Minnesota, Minneapolis, Minnesota.
B. R. BUCKINGHAM, 33 Chatham Circle, Wellesley Hills, Massachusetts.
EDGAR A. DOLL, Director Emeritus, Devereux Schools, Devon, Pennsylvania.
HARL R. DOUGLASS, Director Emeritus, School of Education, University of Colorado, Boulder, Colorado.
E. DUNCAN GRIZZELL, Dean Emeritus, University of Pennsylvania, 640 Maxwelton Court, Lexington, Ky.
ERNEST HORN, Professor Emeritus of Education, State University of Iowa, W307 East Hall, Iowa City, Iowa.
EARL HUDELSON, Professor Emeritus of Education, West Virginia University, Morgantown, West Virginia.
A. M. JORDAN, Professor Emeritus, University of North Carolina, Chapel Hill, North Carolina.
ALFRED S. LEWERENZ, Director Emeritus of Evaluation and Research, 4107 Holly Knoll Drive, Los Angeles, California.
A. R. MEAD, Director Emeritus of Educational Research, University of Florida, 1719 N. W. 6th Ave., Gainesville, Florida.
M. J. NELSON, Dean Emeritus, Iowa State Teachers College, 6033 West Orange Drive, Glendale, Arizona.
JESSE B. SEARS, Professor Emeritus, Stanford University, Stanford, California.
CARLETON W. WASHBURNE, Chairman Emeritus, Department of Education, Brooklyn College, Brooklyn, New York.

Staff

JEAN WILLIAMS, Supervisor of Editorial Services. **CLARENCE A. SCHOENFELD**, Editorial Associate. **ARNOLD CAUCUTT**, Business Manager.

JeR

The Journal of Educational Research

Dedicated to the Scientific Study of Education

VOLUME 56 **SEPTEMBER 1962** **NUMBER 1**

It Says in The JER 2

Among Those Present 4

Some Issues in Beginning Reading Instruction—Theodore L. Harris 5

Attitudes and Beliefs as Sources of Vocational Preference—A. G. Sorenson and Irma E. Morris 20

The Comparative Effects of Small-Group vs. Teacher-Led Discussion Sessions upon Student Achievement and Perception in Educational Psychology—William J. Gnagey 28

Do Parents Help Gifted Children Read?—Dan Cappa and Delwyn G. Schubert 33

Scholastic Aptitudes, Vocational Interests, and Personality Characteristics of Journalism Students—Edmond Marks, John D. Vairo, and Martin L. Zeigler 37

Case Studies in the Changing Roles and Statuses of College Teachers—James Moss 41

The Effect of Musical Distraction upon the Performance Efficiency of Children—Alvin Engel and Cynthia Engel 45

Parent–Teacher Relationships in the Early Nineteenth Century—Margaret R. Meyer 48

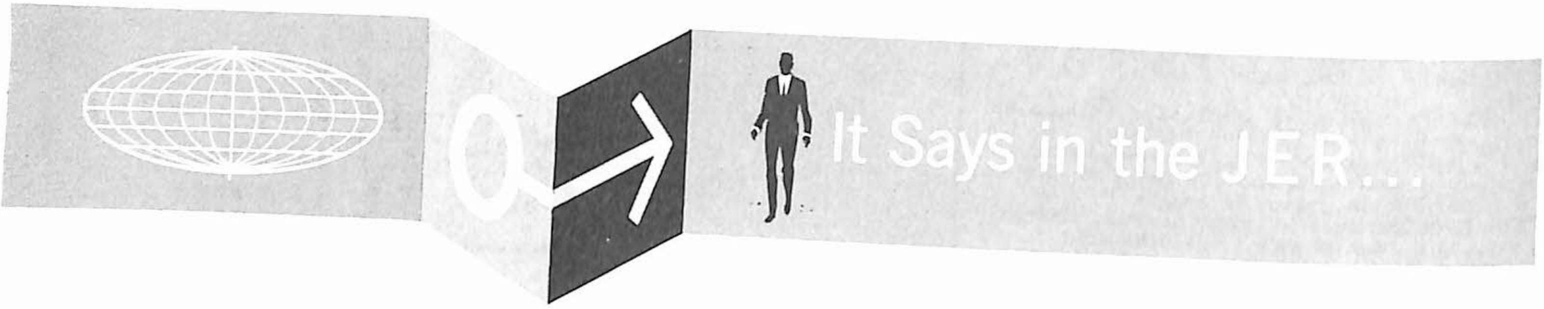
Effectiveness of Different Methods of Study—Mabel S. Noall 51

The Validity of Early Entrance into Kindergarten—Rosalee G. Weiss 53

Selected Characteristics of School Board Members and Their Attitudes Toward Criticisms of the Public Schools—Frank R. Albert, Jr. 55

The Journal of Educational Research is published at Madison, Wisconsin, monthly 10 times a year (the May–June and July–August issues being combined) by Dembar Publications, Box 1605, Madison 3, Wisconsin. The subscription price is $7.50 a year plus $1 a year postage for all subscriptions outside the continental United States. Single copies, $1. Address all communications to Dembar. Second class postage paid at Madison, Wisconsin.

COPYRIGHT 1962 BY DEMBAR PUBLICATIONS, MADISON 3, WISCONSIN


It Says in the J E R . . .

THE P. T. A. MOVEMENT GOES BACK at least to the 1830's, Meyer finds, even though the first Congress of Mothers was not called until 1897. When he called his first meeting in Connecticut in 1839 to “further the cause of the common schools,” Barnard held as part of his program a series of evening lectures for parents. This idea of an evening meeting for the public persisted as a part of the teachers' institutes springing up throughout the East before the Civil War. Already there was the feeling that, “if not the parent, who shall support the suppliant hand of the teacher. . . . in the conflict with besieging hosts which distract the child.”

WHICH IS MORE EFFECTIVE, small-group or teacher-led discussion sessions? Testing college sophomores in educational psychology classes, Gnagey could find no differences between the two methods in terms of student achievement, but the small-group-discussion students tended to rate the course “more practical” and “more understandable” than did the teacher-led-discussion students.

CAN YOU RUSH A CHILD INTO KINDERGARTEN TOO FAST? Yes, says Weiss. Her studies of high-IQ early-age children placed in school with normal-age kindergarteners suggest that such pupils may be “exposed to experiences that may hinder rather than help their intellectual, emotional, and social development.”

BACKGROUND MUSIC DOESN'T DISTRACT pupils in the performance of arithmetic problems, the Engels find, and they go on to suggest that appreciation of fine music could be taught in subliminal fashion.

JOURNALISM STUDENTS who persist in the major exhibit no aptitude, interest, or personality traits significantly different from those who drop out, according to a Pennsylvania state study.

ATTITUDES AND BELIEFS are undoubtedly a source of vocational preference on the part of ninth-graders, but these attitudes and beliefs are based as much on fancy and misunderstanding as on accurate information, Sorenson and Morris report. For example, asked what jobs they "wouldn't like," 130 California students rated teacher, doctor, scientist, garbage collector, and secretary lower than janitor, mechanic, bus driver, policeman, and artist.

WHICH METHOD OF STUDY contributes most to student learning and retention of facts: reading with note-taking, reading, listening with note-taking, or listening? Testing both high-school and college students, Noall finds that on immediate recall all four methods seem equally good, but on delayed recall significant differences appear. The order of superiority is the order listed in the question.

PARENTS HELP GIFTED CHILDREN READ, Cappa and Schubert report. The majority of them are buying books and subscribing to magazines for their children. They are providing time and place at home where free reading can occur. They are critical of their children's reading and are interested in raising its quantity and quality, but they do not level any criticisms against the school.

Among Those Present

JAMES A. MOSS is a sociology professor on the staff of Union College, Schenectady, New York. He holds degrees from Fordham and Columbia. Dr. Moss has taught at the University of Puerto Rico, Dillard, Orange County Community College, Queens College, and the Salzburg (Austria) Seminar. He is currently on leave as director of research for the Southern Regional Council, Atlanta, Georgia. A dozen of his articles have appeared in leading professional journals.

MARGARET R. MEYER has been librarian at Russel Sage College since 1940. Her degrees are from University College at Albany and Columbia. She is the author of numerous professional articles.

ROSALEE G. WEISS is the senior clinical research psychologist at the New Jersey State Diagnostic Center in Menlo Park. Her degrees are from Western Reserve and New York. Dr. Weiss has also served the state of New York and Fairleigh Dickinson University. She is co-author of The Pedophiliac Offender, and received the New York University Founders Day Achievement Award in 1960.

WILLIAM J. GNAGNEY is now associate professor of psychology at Illinois State Normal after having taught in Detroit and at Manchester and Butler. He holds degrees from Wayne State. He is collaborating at present in a comprehensive research project on "Forces Responsible for Shifts in Professional Attitudes of Students in Teacher Education."

FRANK R. ALBERT, Jr., attended Memphis State University where he received his bachelor's and master's degrees, and from the University of Mississippi his Ph. D. He went to the Inter American University of Puerto Rico in 1959 as director of the Educational and Laboratory Studies Division. At present he is associate dean of Co-Campus Teacher Education and director of the Evening Program. He has had several articles published in professional journals, and has memberships in the American Association of University Professors, American Guidance and Personnel Association, Phi Delta Kappa, and the National Education Association.

DELWYN SCHUBERT has been serving as coordinator of primary education and remedial reading for USAF schools in Europe, after being on the staff at Los Angeles State College for ten years. He is the author of The Doctor Eyes the Poor Reader. His PhD is from Northwestern.

THE ENGLES (Alvin and Cynthia) are a man and wife research team. He is currently teaching in East Meadow, New York. She is a psychologist at Levittown schools.

DAN CAPPA is professor of education and head of the elementary education department of Los Angeles State College. He holds a PhD from the University of California. His experience includes that of director of elementary instruction in the Santa Barbara City Schools, as well as 20 years of public school teaching.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 1, September 1962)

Some Issues in Beginning Reading Instruction

THEODORE L. HARRIS
University of Wisconsin

MANY ARE the symptoms of discontent with the effectiveness of beginning reading programs today. There is the parents' shocking dismay in discovering that an able son, well into first grade, still guesses wildly at words and fails to recognize the word he has just "read" on the previous page; or their despair at being told that their third-grade child has done "nicely" in reading but will have trouble in the next grade because he "doesn't know his sounds very well." There is the deep concern of teachers, administrators and remedial reading specialists who have for many years watched the steady procession through the grades of significantly underachieving readers, estimated to be from 10 to 25 percent of the school population by the end of sixth grade (26); and they wonder if educational methods and materials are significant factors in reading underachievement, and if the massive and costly remedial reading programs to which cities[1]* are turning are the only solution to the problem. There are the caustic observations of school critics who, validly or not, unfavorably compare the effectiveness of our reading programs with those of other countries (25), or who champion the presumed superiority of any early "phonic" approach over current "sight-word" approaches (28, 29). There is the strongly-voiced conviction of the linguistic scholar that the common practice of using word frequency and familiarity as the primary bases for word-selection control in children's beginning readers should be replaced by a logical system of letter-sound, word and sentence patterns derived from linguistic analyses of oral written language (4, 12, 19, 23).

These symptoms persist in spite of evidence cited by Worcester and Kline (31) and more recently by Gates (9) that children today as a whole read as well, if not somewhat better, than did their counterparts of a quarter of a century ago. Such a finding is not surprising even for an expanding population in view of the very real recent advances in instructional resources, the awakened concern for individual differences among children, and the heightened sensitivity of teachers, administrators, and laymen to the crucial role of reading at all levels of the curriculum. Yet serious questions symptomatic of public and professional unrest continue to be raised: Cannot more children read even better? Cannot the proportion of underachievers in reading be reduced? Cannot more children be helped to acquire earlier independence and efficiency in basic reading processes to their advantage in subsequent reading tasks?

What are the issues behind these symptoms? Phonics, that popular but much maligned word, is not per se a real issue in reading instruction today. All systematic approaches to the teaching of reading utilize phonics in some way, at some point, and to some extent, —as recent reports sponsored by the Carnegie Foundation illustrate (2, 18). Neither the question of phonics versus no phonics, nor that of the goal of reading—attaining and using meaning conveyed through written symbols—is a significant issue. All reading programs seek to open up the world of ideas expressed in writing. Wherein, then, lie the real issues which contribute to the current public and professional concern? The real issues arise from different conceptions of the nature of the reading process itself and of the learning processes, sets and principles to be stressed which are found in different beginning reading programs. As these differences in thinking about the nature of reading and learning are rooted in strongly held convictions and values, these in turn have led to the setting up of different instructional priorities in given reading programs. Such priorities directly affect the starting point of instruction and the route the child is to take in learning to read.

Two key priorities that distinguish beginning reading programs are those given to meaning, the ultimate goal of reading, and to discrimination, the means to this end—a means with which phonics has been prominently identified. Beginning reading programs can and do reflect very different notions about the route to significant and efficient progress in learning to read in the priorities given to meaning and discrimination. But the current beginning reading controversy cannot be clarified until the specific routes and means proposed by reading programs are related to the different conceptions of the nature of reading and the learning processes involved.

---

*All footnotes will be found at end of article.

Differing Views About Reading

How do conceptions of the nature of reading differ in the way they affect initial reading instruction? One view holds that reading is essentially a cognitive process, of coming to grips immediately and directly with the thought units in our language—the word, the phrase, the sentence, and other more complex patterns of thought. This view, held by Witty (30) and others, represents a meaning-first emphasis which has been the predominant pattern of beginning reading instruction in America in recent years. A second view holds that beginning reading is essentially a perceptual process, one of learning to make the necessary discriminations which our written language requires, especially in relation to our spoken language. This discrimination-first emphasis, one version of which is clearly stated by Bloomfield (4), is an older pattern which lost favor but is re-emerging in various guises especially since the advent in 1955 of Flesch's Why Johnny Can't Read (8) and, more recently, of Bloomfield and Barnhart's Let's Read (5).

Both views accept and work with the functions of meaning and discrimination but assign unmistakably different emphases to these functions at different points on the route to learning to read. As in many issues in which values are strongly held and feelings run high, positions on the teaching of beginning reading have tended to become polarized at two extremes: either emphasize meaning first, then discrimination; or emphasize discrimination first, then meaning. There appears to be singularly little recognition, in spite of protestations to the contrary, that these either-or positions do in fact represent a false dichotomy: the demands of reading upon child or adult and the expectations which readers bring to the printed page require that a delicate meaning-discrimination balance be maintained from the very inception of reading instruction.[2] A third view, therefore, that initial reading instruction should represent a balanced meaning-discrimination emphasis appears to be highly tenable and desirable. Each of these three views will be examined in relation to actual instructional practices after noting the different learning emphases which characterize the meaning-first and the discrimination-first approaches.

Differing Learning Emphases in Beginning Reading Instruction

What assumptions do these approaches make about learning processes, principles, and sets to be stressed in beginning reading instruction? In the initial steps in the meaning-first approach, the assumption is made that associations may be established directly between the sight of the written whole-word form and its meaning (30). In this view of verbal learning, the written word serves as a sign of an idea with which the child is familiar and has had oral language experience. Associations between the sight, sound, and experience relevant to a word establish its meaning and recognizability. That learning may take place in this manner in human beings is commonly assumed in associative theories of learning.

The principle of readiness is particularly important in the meaning-first approach in order to insure that the child has the appropriate language and experience background to bring meaning to the words he is asked to identify. Word recognition is encouraged by control of word frequency and familiarity, and retention is stimulated by a high degree of controlled repetition of words and by elaboration of the meaning of these words in a variety of contexts to strengthen the sight-sound-meaning associations. Transfer of learning in initial reading experiences presumably takes place chiefly in the form of a generalized attitude toward the importance of meaning in reading. An especially critical learning principle invoked in this approach is that of immediate motivational satisfaction (18). Encouraging a child at the outset to attempt and to succeed in reading a simple story is assumed to provide the child with satisfaction, the attainment of the story meaning thus providing its own reward. In its concern for the cognitive or meaning aspects of learning, directly reflected in the primacy given to word wholes and word patterns as the basic carriers of meaning, this approach reflects one influence of Gestalt psychology (6).

In the discrimination-first approach, an associative process is likewise assumed to take place. The associations emphasized, however, in the first step of this approach are those between the written symbols of our alphabetic language and their respective sounds in speech, not between whole-word forms and the ideas to which they refer. (Illustrative rationales for the discrimination-first approach may be found in Diack (6), Bloomfield (4), and Hall (12).) It is then assumed that the discriminations thus developed may be used by the child as a second step to unlock the sounds of words and, if familiar, determine their meaning. Repetition is likewise prominent in this approach but is directed principally to practice upon controlled symbol-sound patterns rather than to control based upon word and meaning frequency counts. Transfer of the early learning of discriminations in this approach is probably a function of the extent to which the discriminations are generalized. Readiness in the sense of familiarity with the spoken language is particularly important, although one of the chief functions of this approach is to build specific readiness in symbol sight-sound discrimination for the later unlocking of new words independently. Motivational satisfaction in the initial steps of this approach is derived chiefly from the sense of independence and achievement the child acquires from learning to make discriminations among words rather than from the satisfaction of immediately reading a story.

The Significance of These Differences

One is tempted to say that the most significant comment upon these differences is the fact that they exist. Faced with the obvious difficulties of building a reading program to meet the needs and characteristics of young children on the one hand, and the admitted complexities of the English language on the other, we have tended to resolve this dilemma in one of two ways: either by emphasizing initially the goal of reading as the attainment of meaning, or the means of reading, discrimination, as the essential first step in learning to read.

In effect, these approaches are complementary. Emphasis upon meaning in the reading act is the strength of one; emphasis upon efficiency in symbol discrimination is the strength of the other. The demonstrated weakness of the one is that it frequently fails to provide early and sufficient discriminations to support the emphasis upon whole words and upon meaning. As a result, far too many children come to rely upon memory-association techniques with obvious accompaniments of frustration or boredom as they face a highly repetitive, rigidly-controlled vocabulary of short, "easy" yet hard-to-discriminate words. On the other hand, early overemphasis upon discrimination of parts of words as the sole objective of a beginning reading method may readily preclude proper attention to the role of meaning in the reading act not only for words but for the meaning patterns of phrases, sentences, and paragraphs. In such instances, the repetitive contexts of the whole-word approach may be replaced by equally restrictive isolated drills upon symbol-sound associations as ends in themselves with little application to reading experiences that are meaningful to the child.

It is evident, too, that the meaning-first and discrimination-first approaches are designed to produce very different learning sets in young readers.[3] The meaning-first approach initially develops a set toward whole-word perception which must then be modified as part-word discrimination training is introduced. While many children succeed in making this transition, many others become so conditioned to seeing words as unstructured whole forms that attempts to alter this set are frequently difficult and ineffective. Contrariwise, the child with a strongly-oriented discrimination set may become so habituated to examining the details of words that he fails to develop a modification of this set to permit him to attend to word wholes and other larger units of meaning necessary for effective fluent reading.

An analysis of specific instructional practices representative of the meaning-first and and discrimination-first approaches will further clarify the nature and significance of these differences in these approaches.

Whole-Word, Meaning-First Practices

In a typical first pre-primer reading activity, a child opens his book to find a word such as "Tom" below the picture of a boy. The teacher may use a number of ways to introduce this word. She may have children discuss the picture and speculate as to what the boy's name might be, call attention to the word's form and sound, and ask for ways in which this word might be used in sentences. Basically, however, the child is told what the word is, is asked to study what the word looks and sounds like, and is helped to remember that the name of this boy is "Tom." This is necessary because the child initially has no equipment to learn the word other than by arbitrary association of the sight and sound of the whole word with the picture clues in the reader and the spoken clues given by the teacher. He will meet this word many times thereafter, and new words will be introduced at a carefully controlled rate since it is known that, on the average, several dozen repetitions are needed for most children to read by a whole-word approach.

In this way, the child is gradually introduced to a sizable number of whole or "sight" words to build up a basic stock of words for ready recognition, presumably by cues of word form and by the context in which they are used. Vocabulary control is primarily on the basis of word frequency and familiarity. Furthermore, since it is extremely important to give children immediate success and pleasure in learning to read something, single words are soon grouped into phrases and sentences to provide the satisfaction of reading a story. If the child experiences difficulty in recalling a word as he is reading a story, the teacher usually gives him a broad hint or tells him the word, thus attempting to keep the reading process, and hence the motivation, alive.

After the child has acquired a substantial base of sight words, a particular type of phonics[4], emphasizing initial consonants, is typically introduced first. A principle of substitution of sounds is used to explain, for example, why words which are similar in form except for the beginning consonant are pronounced differently. If the child knows the word "cat" and now encounters "hat," he will be shown that the "-at" is constant in form and sound, but that the beginning letters look different and sound different. The same substitution principle may now be extended to cover other words in the "-at" family, such as "mat," "fat," and "rat," and to other groups such as "-in," "-an," and "-it" words. Differences among final consonant sounds, and lastly, differences among vowel sounds, are then similarly taught by the substitution principle. Attention to phonetic generalizations, to principles of syllabication, and to the analysis of words into such components as prefixes, roots, and suffixes is usually considerably delayed in this approach. Because the focus in learning to read is first directed toward word wholes and only later to their parts, this approach is often called the analytic approach.

Word-Part, Discrimination-First Practices

In one example (13) of this approach, the child is first given training in the seeing and hearing of the basic short vowel sounds. A few consonants are next introduced. The child then learns to build short words by varying vowel sounds within given beginning and final consonants. For example, the initial consonant b plus a vowel plus the final consonant d might be used to make such words as bad, bed, bid, and bud. This process is then extended to other letters and letter combinations and to words which have some similarities in form and sound. Especially contrived contexts are used to give specific practice in establishing the word groupings thus introduced, some being much more artificial than the use of the sentence "This is a bad bed for Bud," to illustrate three of the four words above.

In this approach, early and concentrated attention is given, again by a principle of substitution, to likenesses and differences in the sounds of similar word forms. The substitution principle, however, is applied differently than in the whole-word approach, in that the child is immediately introduced to several vowels around which consonants are built to form words. Principles of syllabication and the analysis of words into their structural parts are introduced relatively early. Since word parts are first learned and then used to build words, this approach is commonly called a synthetic approach.

A variant of the word-part approach, the "phonemic" or unit-sound approach suggested by Bloomfield and described in the recent Bloomfield and Barnhart's Let's Read (5), starts with the basic short vowel and consonant sounds which the child must first learn. Short three-letter words which have been established by linguistic analysis as being phonetically regular, are next introduced by word families in the approximate order of their frequency of occurrence in the English language. For example, the "-an" family, which is treated as a word part, is introduced first in the form of such words as can, Dan, fan, and Nan--again by a principle of substitution. Sentence drills giving intensive practice in these words follow:

Nan can fan Dan.
Dan can fan Nan.
Dan, fan Nan.

Can Dan fan Nan?
Nan, fan Dan.
(5:60)

Two principles are invoked to determine the progression of such practice drills: 1) to move from the phonetically regular to the phonetically irregular alphabetic spellings, and 2) to move from simple to complex word forms. It is the use of such stringently controlled progressions based upon scientific analyses of spelling-sound frequencies that chiefly distinguish this approach from "phonic" approaches that are approximately a century old. The analogy here is one of coding: Learn the basic spelling-sound patterns first and then apply these skills later to what Bloomfield calls "ideational reading" (4). While the "run, run, run" contexts of word-meaning type readers have been the butt of many criticisms and jokes, the reader will have to decide if the highly artificial and contrived contexts of such a discrimination-first approach represent an improvement over those of the meaning-first approach in the idiomatic use of language and in the appropriateness of its use with children.

The Instructional Dilemma

The contrasting instructional approaches to beginning reading which have been described present in each case an instructional dilemma which needs to be resolved. Many questions may be raised in considering a desirable resolution of this dilemma. Is the meaning-discrimination imbalance reflected in these initial reading programs defensible? What is the basic discrimination unit that should be emphasized in beginning reading—the whole word, the word part, or some larger whole? Can rote associative learning characteristic of both approaches be avoided in the early stages of learning to read? What, other than experiential familiarity, should serve as the basic control on vocabularies in readers—word frequency or some type of linguistic regularity? How economical and transferable is the substitution method of teaching word analysis skills? Should the principle of immediate motivational satisfaction of reading a story in a pre-primer take precedence over satisfaction in developing independent word perception skills? And what model do we have of the child as a learner—that of a pseudo-machine to be activated through stereotyped drills, or that of a thinking organism with developing perceptual and cognitive powers of discrimination and meaningful generalization?

Can the dilemma posed by these approaches be resolved? A truly effective reading program should manifestly emphasize the strengths of these approaches and avoid their respective weaknesses. It should endeavor to make the search for meaning active and continuous in the reading process from the very beginning of formal reading instruction. It should also build critical discrimination skills from the outset so that they are of maximum value throughout the process of learning to read. Such a reading program in other words should seek a new meaning-discrimination balance to resolve the current dilemma posed by the imbalance in the meaning-first and the discrimination-first approaches.

A Dual Meaning-Discrimination Approach

Examination of a new pattern of beginning reading instruction will prove instructive in consideration of the questions raised in connection with the approaches previously described. An analysis of a primary reading program which has been widely used in this country since 1952 shows that it differs significantly from that of either the meaning-first or the discrimination-first approach in several respects: 1) in the nature of beginning experiences in meaningful reading, 2) in the sequence, timing, and emphasis developed in discrimination activities, 3) in the key learning processes, sets, and principles emphasized, 4) in the time of introduction of independent reading, and 5) in the way in which the meaning-discrimination balance is maintained in the independent reading of new materials. Since a comprehensive analysis of the rationale of the dual meaning-discrimination approach represented by this reading program has not appeared in print, the instructional procedures and learning emphases of of Phonetic Keys to Reading (22), a basic reading program for grades 1, 2, and 3 will be described in some detail, and the effectiveness of this program appraised.

The Meaning-Discrimination Balance in Initial Reading Activities

The initial reading activities in this program are of two types: those designed to help the child understand the way in which written symbols are used to convey meaning, and those used to build the basic discriminations needed for early independence in reading. Parallel emphases upon these two types of activities at the very outset of formal reading instruction represents a new balance between the functions of meaning and discrimination in reading. These activities are introduced prior to having a child attempt to read independently a story in a pre-primer.

If the active seeking of meaning through the use of whole words and sentences early in the reading process is to be a realistic goal of beginning reading instruction, the child should be provided with experiences in the reading of symbols that have maximum meaning for him. The conditions under which he actively, rather than passively comes to understand how written symbols are used to express ideas are likewise very clear: such understanding most appropriately comes from the child's own language as he describes a meaningful experience of his own. As he sees, for example, the beginning words of his own story, "I have a white rabbit," put on the board by his teacher, each symbol and the whole pattern of symbols are immediately full of meaning for him. The psychological significance of the act of attributing meaning to symbols under these conditions should not be confused with the problem facing the beginning reader who for the first time sees these same words embellished with a suitable illustration in a reader. The child may well conclude because of the illustrative context that one of these symbols must be "rabbit"— though it could be "hare", if he is familiar with Aesop's "The Tortoise and the Hare". But if he lacks the basic perceptual skills in reading, he is unequivocally put in a position of guessing and of almost complete dependence upon the teacher or an illustration for the clues to the meaning of each word in a sentence.

The child will come to understand the nature and function of reading as a process most actively and meaningfully in initial reading activities if the symbols he is expected to read come from his own experience and represent his own idiomatic use of language. Experience chart work in this way provides the pupil with models of sentence patterns with which he is familiar, with experience in the form and sequence of whole words and phrases, and with practive in the left-to-right orientation required for reading the English language. Here, then, is a clear alternative in an initial reading-for-meaning activity to that of asking the child to begin to read stories via a sight-word approach before he has the skills to make the process of reading truly active and meaningful.

The early emphasis upon meaningful reading in this program is paralleled by work in a separate period in building basic perceptual skills in reading — the phonetic relationships between letters, spellings, and their sounds— to prepare the child to enter into the independent reading of new words and stories actively, meaningfully, and at an early date.

How Phonetic Relationships are Taught and Learned

Unlike the beginning sound substitution practices characteristically found in meaning-first or discrimination-first approaches, phonetic relationships are taught in the dual meaning-discrimination approach as letter-sound discriminations to be identified in a great variety of word forms and meanings. Indeed, the development of any broadly applicable, meaningful relationship or generalization requires that the necessary discriminations be clearly identified and their presence noted in many different situations. In the dual meaning-discrimination approach, this principle is followed in teaching the specific sight-sound relationships of vowels, consonants, and special spellings as well as in developing the more comprehensive phonetic generalizations relating to classes of sounds.

The child is introduced first to the long sounds of vowels, one at a time. The short vowel sounds are then taught. The vowels are taught first because all words and syllables contain vowels, and because most consonants involve vowel sounds in their pronunciation and cannot properly be taught as isolated sounds. The latter point has been noted by linguistic scholars such as Bloomfield (3:501), widespread instructional practices in reading to the contrary. The chief consonant forms are next identified visually and their sounds explored in relation to different vowels, thus emphasizing the importance of consonant-vowel blends.

The introduction of a new vowel such as the long i illustrates how a specific type of discrimination is developed in this approach through varied activities with words in which this letter-sound appears. Such activities include identifying the long i as a distinctive letter form which some children may volunteer to name, listening for this sound in naming familiar pictured things such as knife, mice, and pipe, and marking the long i in other familiar words such as dime, side, ripe, mine, and like, which are placed on the blackboard. Word and story meaning emphasis is introduced by showing the children five additional words pictured in the pre-primer—line, fire, pie, vine, ice— and by asking them to listen for the these words as the teacher uses them in the context of an interesting story told to the class.

Discussion of the story meaning follows, after which the specific words containing the long i sound in the story are identified in the pre-primer, and the distinguishing letter i marked. Additional experience is given also in identifying the long i in other words shown at the bottom of the pre-primer page: hide, five, smile, tie, and bite. In all, the children examine the sight and sound of the long i as it appears in some 18 different word forms exclusive of those volunteered by the class. The significant visual and phonetic component common to each of these words is thus experienced widely. Specific attention to the meanings of key words and to the story in which they are introduced provides a functional situation for relating discrimination and meaning in learning to read.

Through this approach, the child learns as a part of his beginning activities in reading to perceive words as differentiated wholes. He is encouraged to look for and find critical likenesses and differences in a carefully selected group of phonetically related words. He is given help in establishing the habit of active scrutiny of word wholes and parts in a search for their sound and meaning equivalents. He is aided in discovering phonetic relationships within dissimilar as well as similar word forms. Thus the child learns, before he begins the independent reading of stories, to become selective in his word perception processes in a way that is usually characteristic of only the more mature reader (27: 118-9).

In terms of the earlier discussion of teaching phonetic analysis skills by the word-family substitution method typically found in the meaning-first and discrimination-first approaches, children of course learn some limited phonetic relationships. These relationships, however, are initially sharply restricted to the particular spelling families involved, such as the -at, -in, -an, -it families, or the bad, bed, bid, bud group. Consequently, what the child has learned about cat, tin, ran bit, bad, or bed will not help him learn new words such as fun, cap, tip, mud, not, or am, if what he has learned is confined to the particular spelling-sound families involved, and he fails to see any common phonetic relationship among each of these words.

That there is a basic phonetic generalization governing short vowel sounds which cuts across the word-family groups illustrated in the preceding paragraph is very obvious. At a very early point, the child therefore learns in the dual meaning-discrimination approach two types of simple, yet powerful phonetic generalizations: the conditions under which vowels are usually short or usually long in words of one syllable. By knowing the chief vowel and consonant forms and their sounds, and by working with these basic phonetic generalizations, children are able to unlock the key vowel sounds and link appropriate consonant-vowel sounds because they know the principle governing vowel sounds in such diverse words as see, ride, take, goes, clean; me, so, fly; or can, this, just, best, got.[5]

There are, of course, many exceptions to such phonetic generalizations or rules. The problem here is one of the basic complexities of the English language which any approach to the teaching of reading must face at some point. This is precisely why three guiding instructional principles of the dual meaning-discrimination approach should be explicitly noted: a) to control the phonetic complexity of words to which the child is first introduced; b) to teach the child always to use contextual meaning as a referent in checking upon word perception accuracy; and c) to carefully control the rate of introduction of new discriminations and generalizations.

As in any properly graduated learning task, new principles are introduced only as earlier ones become habituated and automatic. Under these conditions, the child readily learns new phonetic generalizations and applies them to progressively more complex spelling-sound relationships. Helping the child see and use these important relationships enlists the most important human ability we know--the ability to make significant generalizations. Failure to do so at the earliest opportunity, as phonetic substitution approaches customarily do, raises serious questions about the efficiency and meaningfulness to the student of the latter type of phonics instruction.

Introduction to Independent Reading of Stories

When the child is next introduced to the independent reading of stories in the pre-primer, he is expected to carry over or transfer his developing discrimination skills and phonetic understandings to discover "on his own" the meaning of a story he has never seen before. Here, then, is an extremely important distinction in the nature of the expectancies of the child and the teacher at this point in the dual meaning-discrimination approach. If a pupil is asked to read, for example, the sentence, "Pat and Mike ran home," he would now attempt to put to work his acquaintance with sentence patterns derived from his experience chart activities and the discriminative and generalizing skills learned in his word analysis activities. Similarly, the teacher would expect him to do so and would refain from helping him unless it were clear, for example, that he needed further help in understanding a phonetic principle involved. Thus, while certain portions of the texts of this beginning reading program may appear similar to those which emphasize a whole-word approach, the nature of the pupil and teacher activity with respect to the reading of the stories is entirely different.

With the introduction of independent reading materials, the development and maintenance of word perception skills is continued by specific daily provisions for extending principles of phonetic analysis to new words, for developing new phonetic principles, and for reviewing word perception skills previously learned. The approach to word perception is broadened and the expansion of a broad meaning vocabulary is stressed. The use of context clues to meaning continues to be stressed. Relatively early, the child is introduced to principles of syllabication and structural analysis and is given much experience in reading longer and more complex words. In these ways, the child develops and uses many resources for word analysis which enable him to attack new materials independently and thoughtfully. In so doing, the child is brought into contact with and is helped to develop a much wider reading vocabulary early in the reading program than is characteristic of the whole-word, meaning-first approach.[6]

Separate periods are devoted to the development of word perception and vocabulary-building skills and to the development of understanding, appreciation, and enjoyment of stories through independent reading. The development of a wide meaning vocabulary is stressed in both periods and forms an important link between the two types of activities, the instructional sequence being so arranged as to promote the transfer of growing word perception skills to the reading of stories. The processes of comprehension and interpretation in oral and silent reading play an early and expanding role in reading instruction. Wide independent reading of supplementary materials is encouraged. An emphasis upon important organizational, locational and evaluative skills in reading is made possible because of the early attention given to the basic meaning and discrimination aspects of the reading process.

Learning Emphases in the Dual Meaning-Discrimination Approach

This approach offers a unique resolution of the meaning and discrimination imbalance found in beginning reading programs of the meaning-first or discrimination-first type. It immediately introduces the child to parallel activities, a) in understanding how symbols are used in language patterns to convey ideas under conditions of maximum meaning, the experience chart; and b) in acquiring the perceptual discrimination skills essential for independent reading. It replaces narrow rote associative learning tasks with learning through meaningful discrimination and generalization. It regulates successive learning tasks by employing a principle of phonetic regularity and frequency in vocabulary control that is more liberal than that observed in discrimination-first approaches, yet avoids the extremes of phonetic irregularity characteristic of meaning-first approaches. It develops motivational satisfactions of two types, that of increasing independence and pride in discrimination skills, and that of the increased satisfaction of initially reading stories with a minimal amount of teacher assistance.

Critical readiness skills in discrimination are actively built; not their emergence merely waited for among pupils who possess adequate ability to learn to read. A significant learning principle which clearly differentiates this approach from those previously reviewed is its emphasis upon generalization as a vehicle for optimum transfer of learning. A broad base for generalization is laid in the rich variety of words used to develop phonetic generalizations, and applications of appropriate generalizations is encouraged in new and varied situations. As the relationship between discrimination and meaning becomes more clearly understood by the child, a favorable attitude toward the total reading process emerges.

The learning model which the dual meaning-discrimination approach represents is that of an active learner with capacities for discrimination and generalization which should be tapped and shaped as early as possible to enable him to engage independently in reading as a meaningful, thinking process. It would appear that this approach significantly affects the set of the beginner in reading. The parallel whole-word and part-word emphases teach the child to perceive words as differentiated wholes from the inception of reading instruction. The focus upon early vowel discrimination and generalization is an important feature of this set. The characteristic of active independent attack upon word and sentence meaning is further evidence of a more generalized set produced by the dual meaning-discrimination approach.

The Effectiveness of the Dual Meaning-Discrimination Approach

Two important considerations in judging the effectiveness of any learning approach are what shall be appraised and how shall this be done. In the appraisal of such a complex activity as learning to read, these are difficult questions to answer. Further problems in appraisal arise from the fact that learning to read involves not only pupils and materials but also important teacher-pupil-materials interactions. The comparison of a relatively new approach such as the dual meaning-discrimination approach with a conventional approach may likewise involve strongly-held attitudes among teachers, parents, and experimenters.

If the foregoing analysis of learning emphases in the dual meaning-discrimination approach is valid, pupil changes should be explored in this approach in at least two directions: in the nature of the pupil activity in the reading process and its effects in related curricular tasks, and in relevant learning outcomes produced. The teacher's role and her problems in adjusting to a new approach should be appraised, the attitudes of parents investigated. A variety of appropriate appraisal instruments should be used, including, but not restricted to the use of standardized tests, to ascertain different types of effects. Care should also be taken to account for the representativeness and stability of the findings and to interpret that data accordingly.

To the best of the writer's knowledge, the most thorough, comprehensive, and continuous appraisals of the effectiveness and stability of the dual meaning-discrimination approach are those which have been conducted over a period of ten years by Mrs. Margaret Henderson Greenman, Director of Elementary Education in the Champaign, Illinois, public schools (10,11,14). It should be noted that the Champaign Schools had conducted a superior reading program in the development of vocabulary and comprehension by using a meaning-first approach for many years, but their reading test results showed that the greatest area of weakness was in the teaching of word perception skills. A pilot experimental program using the dual meaning-discrimination approach was introduced gradually and cautiously, after a year of analysis of the new materials, in the basic study from 1952 to 1955, at which latter time a comprehensive and thorough evaluation of the effects of this approach was made.[7]

In the Champaign study, the clear-cut statistical superiority of the dual meaning-discrimination approach over the meaning-first approach in third-grade pupil performance upon three standardized reading tests and one spelling test has been well-publicized (14:23-5).[8] The tests were especially selected because of their emphasis upon vocabulary, comprehension, and work-study skills. All tests were administered by trained examiners, not by teachers. However, the published report also cites five other types of general conclusions, plus ten concluding observations by the investigator, relating to data gathered by observation, parent and teacher questionnaires, discussions with teachers, and the like. Representative findings relating to the effectiveness of this approach in process as well as outcomes are these statements taken from the concluding observations:

2. The children showed almost unbelievable pride in their reading independence. When they were puzzled by an unfamiliar word, they asked, "Help me sound it," not "What is this word?"

3. These children were not confused by the phonetic principles or terms used. Both soon became a part of their own speaking vocabularies.

5. The children were not confused by the letter that had two or more sounds. If they tried one sound and it didn't make sense, they immediately tried another.

9. At the first-grade level the children wrote stories more creatively, and more as they speak, than first-grade children usually do.

10. The children misspelled words but with different types of errors. The word usually looked as it sounded, and it was easy to read and understand their stories.

Among the general conclusions, these four statements may be found:

6. All of the information obtained from parents and teachers showed evidence that pupils' interest in reading was sustained at a high level in the experimental group over the three-year period.
7. Most of the parents and teachers felt that the ability of the experimental group to read more independently also helped in arithmetic, in social studies, in creative writing, and in other uses of language.
8. Most of the parents felt that this new method helped their children in learning to spell. The test results verify this opinion.
9. Observations in the classroom produced no indications that the experimental method produced word-callers or word-by-word readers.

Since these conclusions and the data to support them are most relevant to a comprehensive appraisal of the effectiveness of this approach to beginning reading, it is unfortunate that they have been ignored in published reviews and discussions of this study.

Two other previously unreported effects of the dual approach are noteworthy. An analysis of the original data has shown that there are no significant differences within the experimental group in the number of boys and girls who achieved below grade expectancy. This is in marked contrast to the usual findings in which the number of boys with low reading achievement typically far outnumber the girls. A second effect noted by the investigator has great practical significance for schools, parents and children. True remedial reading problems due to failure to develop word perception skills for educational reasons have virtually disappeared. The investigator has informally estimated that in ten years of experience with this approach, 95 percent of the children who were mentally able to learn to read, have learned to read. The failure of the remaining five percent to learn to read is believed to be due to primary causes of reading disability such as prior conditions of brain damage or emotional instability.

Further study was made of the stability of performance of third-grade Champaign pupils taught by the dual meaning-discrimination approach over a period of several years to determine if the findings of the original study were merely a novelty effect. Mrs. Greenman compiled the data given in Table I to show the results of third-grade testing over a four-year period in which the number of teachers and pupils using this approach increased steadily. Mrs. Greenman wrote:

> This table demonstrates the results for experimental and control groups over the four-year period except for 1958, when there was no longer any control group. While there are minor differences in percentage figures, basically the picture remains the same for the experimental group even though the number of children and teachers steadily increased. The last bar with the results for 1958, when 29 teachers and over 660 pupils were involved, is strikingly similar to the first bar which shows the results obtained in 1955, for five teachers and 88 pupils. There is also a marked similarity in the results obtained for the control group during three different years. This can be accepted as one indication that the experimental teachers were representative of the total faculty. It also suggests that the Hawthorne effect is not marked in this study, for if it were, surely the novelty of the situation would have worn off in the six years this approach has been used. (11)

This is indeed impressive evidence of the effectiveness and the stability of the dual meaning-discrimination approach as it was gradually introduced over a six-year period in this school system. The two reports tell a story of the careful staff planning needed to effect a shift to a new approach to basic reading instruction without distorting the values which every reading program hopes to attain.

One further enlightening footnote in the Champaign series of studies of primary reading is contained in a recent analysis by Mrs. Greenman of the usefulness of standardized reading readiness and intelligence tests in predicting the reading achievement of 938 first-grade pupils taught by the dual meaning-discrimination approach (10:14-5). Comparison of predictions and actual reading achievement in percent by quartiles, according to national norms, are shown in condensed form in Table II. These test results demonstrate quite clearly that the usefulness of such tests in predicting the level of first-grade reading achievement is to a great extent a function of the type of reading approach for which the predictive values of these tests were standardized. A similar finding was also noted in the 1960-61 report of the Primary Reading Program of Washington School District No. 6, Phoenix, Arizona, in a study of over 1000 children who used this approach in grades one and two (20). The same study also reports evidence of improved arithmetic and study performance, greater ability to write independently, wider reading of supplementary materials, and greater willingness "to read materials not related to their school books" among children using this approach.

Such evidence, broadly conceived to appraise many indirect as well as direct effects of the dual meaning-discrimination approach, strongly suggests that performance upon standardized reading tests is strengthened in the primary grades, that work-study habits and attitudes in related curricular tasks is improved, and that the number of pupils reading below expectancy is reduced to a degree not found in the meaning-first approach. The Champaign studies have shown the effects of the dual meaning-discrimination approach to be very stable over a period of years. The presumption is likewise strong that such a primary reading program will have important continuing effects upon later progress in the intermediate grades.

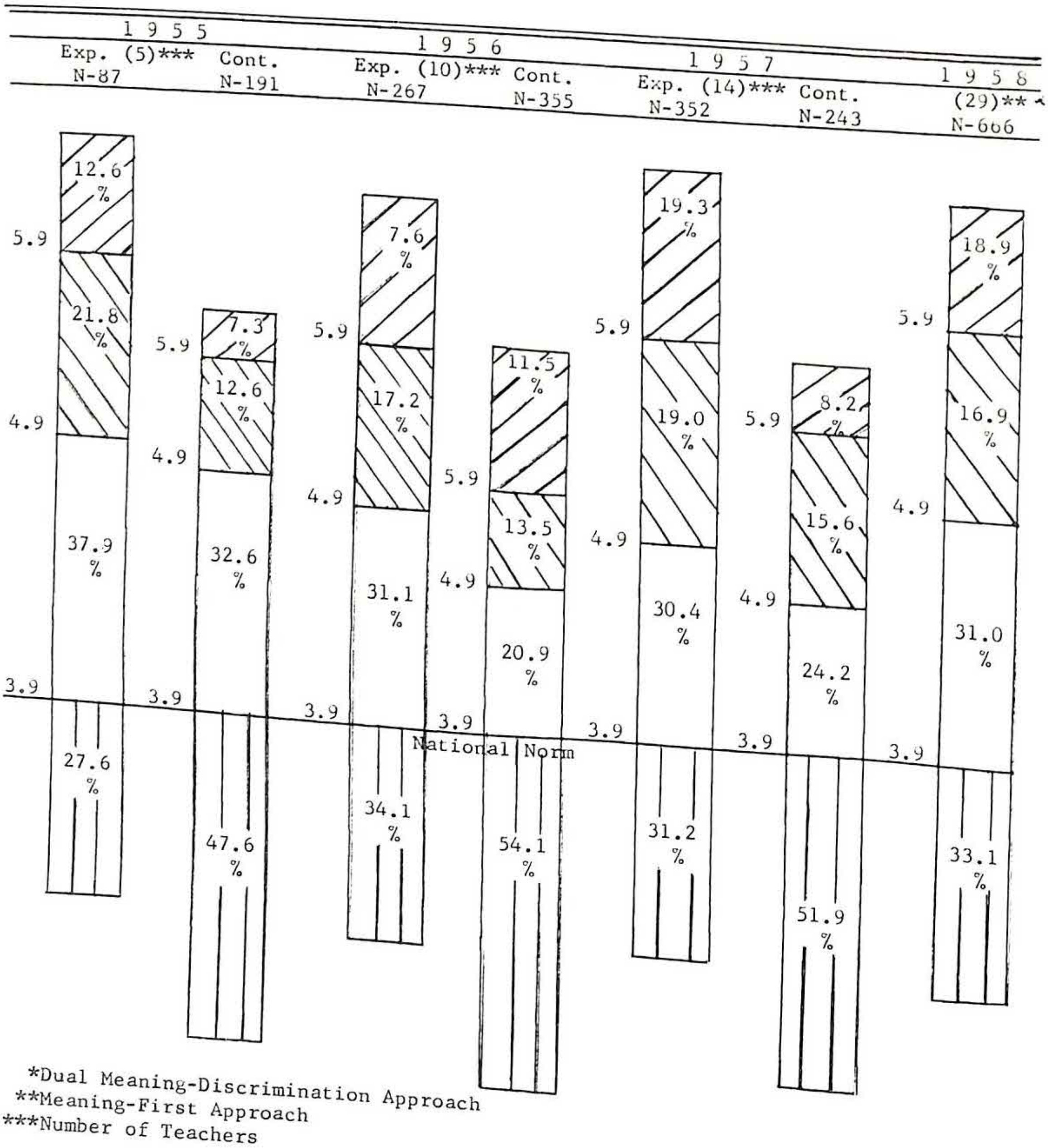
Striking evidence of the holding power of the dual meaning-discrimination approach and its related effects in grades four, five and six is contained in the 1960-61 Report of the Annual Testing Program of the Champaign Public Schools (10:23-5). Comparative median test scores upon the Iowa Tests of Basic Skills are shown for a three-year period for each of these grades in Table III. Each score reflects a group of several hundred pupils. The uniformly superior level of performance of these pupils and their consistent growth from year to year, especially in reading, work-study skills, and the language arts, is noteworthy.

Other evidence of continued reading progress from grades one through six is shown by recent reading test results for 5725 pupils in a metropolitan school district in which this approach has been used for nine years (21). Table IV shows the median total reading grade scores for these pupils at each grade level. The children represented a normal distribution of intellectual abilities, with IQ's ranging from the 60's to the 150's, and a median IQ of 110. Note in the "Difference" column that the middle child at each grade level is over one year (or one grade) ahead of his reading achievement. Note also that this difference is very stable at each grade level for an average of approximately 950 pupils per grade.

These studies, of course, do not provide data on the comparative effects of two approaches to

TABLE I

COMPARISON OF THIRD-GRADE SCORES MADE BY EXPERIMENTAL GROUPS* ON METROPOLITAN READING TESTS: COMPREHENSION OVER A FOUR-YEAR PERIOD (1955-1958) AND THE SCORES MADE BY CONTROL GROUPS**OVER A THREE-YEAR PERIOD (1955-1957)



*Dual Meaning-Discrimination Approach
**Meaning-First Approach
***Number of Teachers

TABLE II

COMPARISONS OF PREDICTION AND ACTUAL READING ACHIEVEMENT

Quartile	Percent Predicted California Test of Mental Maturity	 Metropolitan Reading Readiness Test	Percent Achieved California Reading Test (Total) (May 1961)
	(September 1960)		
Top	24	29	86
Upper middle	32	28	5
Lower middle	26	26	5
Bottom	18	17	4

TABLE III

COMPARISON OF MEDIAN SCORES FOR GRADES FOUR, FIVE, AND SIX FROM THE TEST RESULTS OF THE IOWA TESTS OF BASIC SKILLS FOR 1958-59, 1969-60 1960-61, NATIONAL NORM 4. 1/50th; 5. 1/50th, 6. 1/50th

		Grade Four			Grade Five			Grade Six		
		1958 1959	1959 1960	1960 1961	1958 1959	1959 1960	1960 1961	1958 1959	1959 1960	1960 1961
Reading-Vocabulary	Grade Placement Percentile	4. 5 65th	4. 5 65th	4. 5 65th	5. 5 61st	5. 7 65th	5. 7 64th	6. 9 65th	7. 0 67th	7. 0 67th
Reading-Comprehension	Grade Placement Percentile	4. 5 64th	4. 4 60th	4. 5 64th	5. 5 62nd	5. 5 62nd	5. 7 67th	6. 8 67th	6. 7 65th	6. 7 65th
Language-Spelling	Grade Placement Percentile	4. 5 64th	4. 1 51st	4. 4 60th	5. 3 56th	5. 5 62nd	5. 6 65th	6. 4 59th	6. 5 62nd	6. 7 67th
Language-Total	Grade Placement Percentile	4. 7 69th	4. 3 59th	4. 4 62nd	5. 7 63rd	5. 3 55th	6. 0 69th	6. 4 57th	6. 6 62nd	6. 7 64th
Work-Study Skills Total	Grade Placement Percentile	4. 5 69th	4. 2 58th	4. 3 62nd	5. 3 62nd	5. 4 66th	5. 5 70th	6. 4 64th	6. 5 67th	6. 6 70th
Arithmetic-Concepts	Grade Placement Percentile	4. 6 70th	4. 4 62nd	4. 6 70th	5. 3 61st	5. 3 61st	5. 5 71st	6. 4 66th	6. 4 66th	6. 5 70th
Arithmetic-Problem Solving	Grade Placement Percentile	4. 4 67th	4. 1 52nd	4. 4 67th	5. 0 47th	5. 0 47th	5. 2 57th	6. 3 60th	6. 1 51st	6. 3 60th
Composite-V-R-L-W-A	Grade Placement Percentile	4. 6 70th	4. 3 60th	4. 5 67th	5. 5 54th	5. 5 64th	5. 7 70th	6. 7 67th	6. 6 64th	6. 7 67th

TABLE IV

1960-61 READING ACHIEVEMENT TEST RESULTS IN THE ELEMENTARY SCHOOLS OF THE RITENOUR CONSOLIDATED SCHOOL DISTRICT, OVERLAND, MISSOURI

	Number of Children	Grade Placement at Test Time	Median Total Reading Score	Difference
Grade 1	1097	1. 8	3. 1	plus one year, three months
Grade 2	985	2. 8	4. 3	plus one year, five months
Grade 3	1022	3. 8	5. 5	plus one year, seven months
Grade 4	834	4. 8	6. 4	plus one year, six months
Grade 5	905	5. 8	7. 0	plus one year, two months
Grade 6	942	6. 8	8. 0	plus one year, two months

Grades one and two	California Primary Achievement Tests
Grade three	California Intermediate Achievement Tests
Grades four, five, six	California Elementary Achievement Tests

teaching beginning reading upon later school progress. Furthermore, no data comparable in comprehensiveness and stability to the 1955 Champaign study have been reported to give a precise or conclusive answer to this question. Such evidence as is available is contained in two studies. In one of these, Mrs. Greenman cites follow-up test data on the Metropolitan Reading Test in grades four, five, and six for the experimental and control groups of her original study (11). She reported continued higher mean and median reading achievement test scores, a smaller percentage of pupils below grade expectancy, and greater grade-by-grade consistency in reading progress in the experimental group. While the statistical significance of these differences was not established, the investigator reported these reactions of intermediate teachers working for the first time with children taught by this approach:

> The intermediate teachers who taught these children were very enthusiastic about the work habits these children had developed. They felt that the learning in the content subjects was greatly increased because these pupils could and did use their reading ability to obtain information from all types of reference materials. They felt that they had time to refine the skills of study, do effective dictionary work, particularly for meanings, and to do much word-building study. They agreed that these pupils applied their phonetic ability subconsciously, and that this ability had truly reached the "second nature" learning stage. As one teacher expressed it, "I have time to do the things I have always felt should be done in fifth grade." (11)

A second investigation, the Louisville study by Sparks and Fay (24), comparing the same two approaches in grades one through four upon different reading tests than those used in the Champaign study, showed no significant differences at the end of grade three between the groups in achievement on reading tests, but did show some superiority at the five percent level in favor of the control group in reading accuracy at the end of grade four. Tests were administered by classroom teachers, not by trained examiners. Fay (7) has since reported that no significant differences were later found between the groups at the end of grades five and six.

Neither of these studies offers a completely satisfactory answer to the comparative effects of these approaches upon later school progress. Both are limited in the scope of the appraisal and fail to establish the stability of the findings for other groups, although the three-year intermediate grade data of the 1960-61 Report of the Annual Testing Program in the Champaign Public Schools are strongly persuasive in this connection. The failure of the Sparks and Fay study to demonstrate any significant differences in the two groups by the end of third grade, contrary to the substantial and stable evidence presented here, raises important questions; in particular, questions about the care taken to help those concerned—administrators, teachers, and parents—to understand and plan for the introduction of a new reading approach, and about the wisdom of having teachers rather than trained examiners administer tests in a research study.

A more satisfactory answer to the question of the effects of this approach upon later school progress will be provided as the dual meaning-discrimination approach is extended into a compatible program of reading materials for grades four, five, and six, and a comprehensive appraisal is made of the effectiveness and stability of such a six-year program in conjunction with other reading programs. In the meantime, as intermediate grade teachers learn to adjust their methods, materials, and expectancies to the levels of attainment of pupils taught by this approach—a most important consideration—the presumption of a continued high level of performance in the intermediate grades remains strong.

Toward an Improved Rationale for Beginning Reading

Today, as never before, we are in need of clear thinking about the issues involved in beginning reading instruction. Powerful pressures are at work to bend public opinion, and hence the schools, in one direction or another—often to extremes. New reading materials and programs are being rapidly developed. Will they represent a forward step in resolving the meaning-discrimination dilemma, or merely a retreat under different guises to well-established value-patterns and modes of thinking? Will their claims be paraded under well-worn, "either-or" cliches such as "mechanics vs. meaning" or "whole vs. part" emphases? Will the current myth that "phonics is phonics" continue to be unabashedly fostered as though manifest differences in phonics programs and in the learning assumptions upon which they are based do not exist? Will we witness a perpetuation of the current tendency to take refuge in vague generalities about "composite" and "eclectic" programs, or "components of a good reading program," without any clear indication of precisely how the specific methods proposed fit together into a consistent and psychologically sound general method?

An effective beginning reading program must fulfill two sets of conditions. It must, first of all, meet the meaning-discrimination demands of the reading act in a way that is internally consistent with the complex nature of reading and the significant conditions attending the learning of the skills, concepts, and attitudes involved. Second, it must be adjusted to the nature and characteristics of learning by children, and not represent the mere imposition of drills more appropriate for adults than children. The fulfillment of either one of these conditions is not sufficient; both are necessary.

This article has described a beginning reading approach that resolves the meaning-discrimination dilemma in a unique fashion. It has explained how and why this approach produces effective learning. It has compared it with other approaches in significant ways. It has presented evidence that it consistently produces, over a period of years, earlier and more effective reading comprehension habits and important related skills and attitudes in children than does the predominantly used meaning-first or sight-word approach. This approach does this without distorting the delicately balanced values of meaning and discrimination in reading as meaning-first and discrimination-first approaches characteristically do. Most importantly, this article has shown that the dual meaning-discrimination approach to beginning reading immediately makes the child a truly active participant in the learning process and enlists his capacities for discriminating, generalizing, and applying his understandings of how our language works to the fullest in learning to read.

The implications of such a new approach to the teaching of beginning reading should be considered carefully. For it is only by a re-examination of some of our fundamental learning assumptions and by the development and thorough testing of explicit learning rationales for instruction in reading that there is much prospect for a breakthrough in reading methodology. But if this is done, we may indeed find improved ways of meeting the meaning-discrimination dilemma, of reducing the incidence of early reading failures, and of making beginning reading a process of independent discovery of meaning which will strengthen the total educational process.

REFERENCES

1. Artley, A. Sterl. "Progress Report on the Champaign Reading Study, 1952-55: A Review and Discussion," Elementary English, XXXIV (1957), pp. 102-5.
2. Austin, Mary C. The Torchlighters (Cambridge: Harvard University Press, 1961).
3. Bloomfield, Leonard. Language (New York: Henry Holt and Co., 1933).
4. Bloomfield, Leonard. "Linguistics and Reading," Elementary English Review, XIX (April 1942), pp. 125-30; (May 1942), pp. 183-6.
5. Bloomfield, Leonard and Barnhart, C. L. Let's Read (Detroit: Wayne State University Press, 1961).
6. Diack, Hunter. Reading and the Psychology of Perception (New York: Philosophical Library, 1960).
7. Fay, Leo C. "A Look at Two Approaches to the Teaching of Reading," in Changing Concepts of Reading Instruction. Proceedings of the International Reading Association Conference, VI (1961), pp. 161-3.
8. Flesch, Rudolf. Why Johnny Can't Read (New York: Harpers, 1955).
9. Gates, Arthur I. Reading Attainment in Elementary Schools: 1957 and 1937 (New York: Teachers College, Columbia University, 1961).
10. Greenman, Margaret Henderson. Report of the Annual Testing Program, Elementary Schools, 1960-61 (Champaign Public Schools, 1961).
11. Greenman, Margaret Henderson. A Six-Year Experimental Study of Two Methods of Teaching Reading in the Elementary School, Report to the American Educational Research Association, (February 1959), mimeographed.
12. Hall, Robert A., Jr. Sound and Spelling in English (Philadelphia: Chilton Company, 1961).
13. Hay, Julie and Wingo, C. Reading with Phonics (Philadelphia: J. B. Lippencott Co., 1948).
14. Henderson, Margaret. Progress Report of Reading Study, 1952-55 (Champaign, Illinois: Board of Education, 1955).
15. Herrick, V. E., Harris, T. L., and Rarick, G. L. Perception of Symbols in Skill Learning, Report to the U. S. Office of Education, 1961.
16. Huey, E. B. The Psychology and Pedagogy of Reading (New York: Macmillan Co., 1921).
17. Kelly, Barbara C. "The Economy Method Versus the Scott, Foresman Method of Teaching Second-Grade Reading in the Murphysboro Public Schools," Journal of Educational Research, LI (1958), pp. 465-9.
18. Learning to Read (Princeton, New Jersey: Educational Testing Service, 1962).
19. MacKinnon, A. R. How DO Children Learn to Read? (Toronto: Copp Clark Co., 1957).
20. Primary Reading Program. Report for the 1960-61 School Year (Phoenix, Arizona: Washington School District No. 6, 1961).
21. Reading Achievement Test Results, Elementary Schools (Overland, Missouri: Ritenour Consolidated School District, 1960-61).
22. Sloop, Cornelia B., Garrison, H. E., and Creekmore, M. Phonetic Keys to Reading (Oklahoma City: The Economy Company, 1952).
23. Soffietti, James P. "Why Children Fail to Read: A Linguistic Analysis," Harvard Educational Review, XXV (1955), pp. 63-84.
24. Sparks, P. E. and Fay, L. C. "An Evaluation of Two Methods of Teaching Reading," Elementary School Journal, LVII (1957), 386-90
25. Trace, Arthur S., Jr. What Ivan Knows that Johnny Doesn't (New York: Random House, 1961).
26. Traxler, Arthur E. "Research in Reading in the United States," Journal of Educational Research, XLII (March 1949), 481-99.
27. Vernon, M. D. The Experimental Study of Reading (Cambridge, England: Cambridge University Press, 1931).
28. Walcutt, Charles C., ed. Tomorrow's Illiterates (Boston: Little, Brown and Co., 1961).
29. Walcutt, Charles C. and Terman, S. Reading:

Chaos and Cure (New York: McGraw-Hill, 1958).

30. Witty, Paul. "A Forward Look in Reading," Elementary English, XXXVIII (March 1961), 151-64.

31. Worcester, D. A. and Kline, A. Reading Achievement in Lincoln, Nebraska, Schools 1921 and 1937 (Lincoln, Nebraska: University of Nebraska Teachers College, 1947).

FOOTNOTES

1. The New York Times reported in 1961 that 250 remedial reading teachers were to be added to the staff of New York city schools.
2. The premise that the adult reader reads exclusively by wholes is false, as Huey (16:115-6) and others have pointed out. Confronted by a new word, even the most skillful reader resorts to basic techniques of word analysis to determine its structure, pronunciation, and meaning.
3. The nature and operation of learning sets in controlling perception of a task, in evaluating performance, and in determining speed in a task has been extensively investigated in children's handwriting by V. E. Herrick, T. L. Harris, and G. L. Rarick (15). The research is being extended to the field of reading.
4. Phonics has to do with the phonetic or sound analysis of the written language symbols used in reading. In the application of this type of analysis to the reading of written symbols, phonics borrows from phonetics, the scientific study of language sounds in speech. The distinction is an important one: phonetics deals directly with the sounds of the spoken language; phonics deals with the representation of sounds in the written language. It is thus an audio-visual type of analysis in which accuracy in relating the proper sound to the proper symbol or symbols is necessary. Both sight and sound accuracy are equally important. Systems which employ phonetics for the analysis of the sounds of the written language are thus referred to as "phonics."
5. These words are taken from a list of the 220 most common words, exclusive of nouns, found in whole-word approach primary readers in a count made by Dolch. Of these words, 190 are one-syllable words. Fifty-five percent of the latter words may be explained by three phonetic generalizations governing the vowel sounds of such words and may be learned by most children in first grade.
6. The extent of this difference in initial vocabularies is illustrated by the fact that an average of 325 new words is introduced in five whole-word approach reading series in grade 1, whereas 434 new words are introduced in the first-grade materials of the dual meaning-discrimination approach. In addition, children in the latter approach work with 1825 words in listening and sounding activities in first grade.
7. See A. S. Artley, "Progress Report on the Champaign Reading Study, 1952-55: A Review and Discussion" (1) for certain criticisms of the study such as the unspecified nature of the approach in the control group, and the comparability of teachers in the experimental and control groups, and of the time devoted to teaching reading.

   Mrs. Greenman answered these criticisms in the 1959 paper delivered before the American Educational Research Association: a) the control group "followed the standard practice of developing an early sight vocabulary in conjunction with early independent reading," using familiar basal reading materials; b) "teachers [in the experimental groups] were typical of our faculty in teaching ability, interest, and in years of experience,"; c) "the amount of time spent in basic reading instruction was the same in both groups... All other areas of the curriculum... continued to receive the same amount of time and emphasis." (11)
8. Highly significant differences at the one percent level, favoring the experimental group, were found on ten out of 14 test scores upon the Stanford, Iowa Every-Pupil, and Metropolitan Reading tests (and the Metropolitan Spelling Test), and significant differences at the five percent level obtained for the remaining test scores. The groups were shown to be not significantly different in intelligence. Compare also the similar findings by Kelly (17) for groups of second-grade children matched according to mental ages in a comparison of the effectiveness of the meaning-first and the dual meaning-discrimination approaches.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 1, September 1962)

Attitudes and Beliefs as Sources of Vocational Preference

A. G. SORENSON, University of California

IRMA E. MORRIS, Los Angeles City Schools

WHILE it seems clear that the experiences which are important in determining vocational interests occur during childhood and early adolescence, there is still much to be learned about the specific nature of those experiences and how they influence the occupational inclinations of young people. In order to gain additional information in this problem area, a questionnaire was administered to 130 ninth-grade students of Mark Twain Junior High School, Los Angeles, California. There were 67 girls and 63 boys in the sample, which represented two high, one average, and one low ability English classes, with IQ's ranging from 73 to 132 as demonstrated on the California Test of Mental Maturity.

The items included in the questionnaire were designed to elicit information about the actual work experience of the sample of urban students and their attitudes and beliefs about both the jobs they had held and occupations in general.

Information was sought in the following areas:

1. What was the nature of the work experience?
2. What did the students like and dislike about the jobs they had held?
3. How did the young people define vocational "success"?
4. What were some of their beliefs about how people advance in their jobs?
5. What, in the opinion of the students, were the characteristics of a good job?
6. What criteria did the students use in evaluating possible future occupations for themselves?

Check lists were provided for most student responses, although the boys and girls were invited to write in answers whenever they believed that the choices offered were inadequate.

Results

The Nature of the Work Experience. Of the sample of 130 ninth graders, 84 had worked for pay outside their homes. Of these, 46 were girls and 38 were boys. For the most part, the jobs consisted of baby sitting, paper routes, yardwork, and either housework or janitor work. A larger range of jobs was shown by the boys. In all, the boys listed 46 jobs for themselves; the girls listed 49. These are cited in Chart 1.

Only four of the 84 students with work experience actually later named an occupational interest related to their work experience. One was the boy who had worked in his father's service station; another was the boy who repaired lawnmowers in his uncle's shop; the Sunday School teacher wanted to be a medical missionary; and one girl who had addressed envelopes in her father's office picked "secretary" as a possible career for herself.

What the Students Liked and Disliked About Their Jobs. The students listed more items that they liked about their jobs than complaints. There were 158 checks and write-in statements for items they had liked and 88 objections to their jobs. In both areas the boys registered more comments, including write-ins, than the girls. More write-in comments were added to the unfavorable list than to the favorable, with 11 boys registering complaints as against one write-in favorable response for boys. Six girls added complaints as write-ins; two girls wrote in favorable statements.

The recorded favorable statements are listed in Chart 2.

For boys, the most frequently-marked favorable response was "the pay," with 39 checking this point. Only 18 girls checked this item, with the largest group of girls, 33, checking the response that they enjoyed "the tasks of the job." This was the second most-favored response for boys, with 19 checks. Fourteen of the boys liked working with adults and being treated as an equal; only seven of the girls checked this point. While 12 boys and 14 girls said they liked the jobs because they were learning something for their future careers, as it was mentioned before, only four students later

CHART 1

Jobs	Boys	Girls	Totals
Babysitting		32	32
Paper route	18		18
Yardwork	12		12
Janitor	8		8
Housework		7	7
Saleswork	2	5	7
Office related	1	3	4
Service station	1		1
Bus boy	1		1
Waitress		1	1
Swimming instructor	1		1
Lawnmower repairing	1		1
Sunday School teacher		1	1
Farming	1		1
Totals	46	49	95

CHART 2

I Liked	Boys	Girls	Totals
The pay	39	18	57
Working with adults and being treated as an equal	14	7	21
Learning a trade for my future	12	14	26
The tasks of the job	19	33	52
Other (write-ins)			
Helping father	1		1
The children		2	2
Totals	84	74	158

CHART 3

I Disliked	Boys	Girls	Totals
Always being told what to do	15	7	22
Learning other people made more money for the same work	10	11	21
Having to work with adults		1	1
The long hours	7	12	19
The hard work	6	1	7
Other (write-ins)			
Moving from yard to yard	1		1
Never being told exactly what to do	1		1
The monotony of the job	5		5
The way people on my route were always causing trouble	1		1
The kids who were monsters		5	5
The boss	1		1
The job because it was so dirty	1		1
Being watched so closely	1		1
What I had to do	1		1
Totals	49	39	88

1S70
J 301

named choices of careers for themselves related to their work experiences.

The 88 negative responses contained 18 write-ins. Nine boys and 18 girls had no complaints to record about their employment; and of these, five girls attached special notes stating they liked everything about their jobs.

The objections students had to their jobs are shown in Chart 3.

Opinions About Occupational "Success." In order to discover what ninth graders thought vocational "success" meant, a series of definitions was provided in a check list. In all,282 responses were given, with three write-ins by girls. See Chart 4.

Responses came from 116 out of the sample of 130.

The answers were, on the whole, highly idealistic and in contrast to the reasons students gave for rejecting occupations for themselves (Chart 7). There were no noticeable differences in responses between boys and girls.

How People Are Promoted. When asked if they believed that promotion was possible within the occupation they were considering for a career, 104 of the sample said "yes," or 80 percent. Sixteen, or 12 percent, said "no." Ten students,or 8 percent, did not respond.

Queried as to what they believed constituted the factors that caused people to be promoted, the students seemed to think that rewards come to those people who do superior work and are willing to do more than is required. Only one write-in statement was added to the check list by a girl : "They are neat in their work."

Answers again seemed highly idealistic, with little difference shown between the sexes in responses except for the items: "They have leadership qualities" and "They are always quick to do more than is required of them." Refer to Chart 5.

Attitudes toward Occupations and Adult Work. When asked if they believed people of different age groups would value a job for the same reasons, 43 out of the sample of 130 said "yes." Sixty -three said "no" and 24 did not answer the question.

When asked to evaluate what they considered important aspects of a good job, the students checked 432 items. Because there appeared to be differences in responses between boys and girls that might be culturally-inspired, this section of the questionnaire was replicated a year later with an entire school population of 9A students, 445 in all: 225 boys and 220 girls. The impression gained from the earlier , smaller sample was confirmed in the larger sample. *In the first columns below are the data from the original study. Opposite are the data from the second study. These are shown in Chart 6.*

While the numbers of boys and girls making the responses were not exactly the same, the bases were very close: 59 boys in the first sample as against 58 girls who responded; 225 boys and 220 girls in the sample studied a year later. Security on the job is of the greatest importance to the girls, although in another question on the questionnaire , the girls stated they preferred marriage to a career. Girls also desired more than boys to do creative work and to have a variety of work tasks. They also considered it more important than the boys to have a clean place to work.

On the other hand, boys considered high beginning pay a bigger factor than did the girls. Although responses were few in number for this point, boys also showed more interest than girls in the prestige element: "A job that other people envy."

On one point only did the trend noted in the first study reverse in the larger sample studied a year later: "Routine work that would not be too difficult." The responses in this category were small for both samples. The larger sample also had a smaller proportion of girls who considered a chance to work outdoors a significant point about a good job than did the original sample, in which almost as many girls marked this item as boys.

In the original sample only one student who might be classified as "superior" mentioned "routine work that would not be too difficult." This information was not obtained for the larger sample. The other 18 responses from the original sample came from students who were all in the lowest section or in the average class.

Jobs They Wouldn't Like. The students were asked to list five jobs apiece that they believed they were sure they would not like to perform and to give the reasons they thought existed for their dislikes. Responses to the question were all write-ins, with students listing a total of 70 different occupations. Below will be found a list of occupations most commonly mentioned and the reasons students gave for discarding them in their consideration of possible careers. Items mentioned less than five times do not appear in Chart 7.

While many students answered with jobs they considered beneath their status: "not good enough for me," "too dirty," "pay not high enough," others, it must be mentioned, answered after evaluating their own abilities. Students were inclined to say they hated the subjects involved, although a few students admitted not being adequate in ability to handle the occupation. Some of the rejections were on the basis of what could be considered "teen - age values," lack of information, especially about relative earnings, and resistance to pressures for occupations that today are high status and of a nature that would cause ambitious parents to prod children into them.

While the three careers for women which generally attract the most: nursing, secretarial and teaching, were chosen by girls the most frequently in the listed occupational choices they made at the conclu-

CHART 4

Definitions	Boys	Girls	Totals
Success means enjoying the tasks of the job	40	45	85
Success means earning the respect of one' s fellow workers and employers	39	40	79
Success means getting promoted on the job	26	22	48
Success means earning the highest possible salary for that kind of work	24	20	40
Success means having a prestige job that other people envy	10	8	18
Success means doing as little as possible for the highest pay	4	1	5
Success means doing the best one can within one' s capabilities (write-in)		2	2
Success means gaining satisfaction from helping others (write-in)		1	1
Totals	143	139	282

CHART 5

Responses	Boys	Girls	Totals
They do superior work	48	51	99
They are always quick to do more than is required of them	30	43	73
They have leadership qualities	27	43	70
They take night school courses and try to improve their efficiency	22	27	49
They move in social circles in which they make helpful friendships	13	8	21
They are good at apple-polishing the boss	8	5	13
Other (neatness in work)		1	1
Totals	148	178	326

CHART 6

	Boys		Girls		Totals	
Good chances for promotion	43	122	37	111	80	233
Security on the job	29	131	44	158	73	289
Variety of work tasks	28	42	40	90	68	132
A nice, clean to work	22	113	35	143	57	256
High beginning pay	34	67	18	32	52	99
A chance to work outdoors	21	54	18	26	39	80
A chance to do original, creative work	11	68	21	72	32	140
Routine work that would not be too difficult	13	19	5	38	18	57
A job that other people envy	9	25	4	15	13	40
Totals	210	681	222	684	432	1365

CHART 7

Occupation	Number	Reasons Rejected
Teacher	35	Monotonous Kids don't like them too many brats Don't like school Enough of them Little chance to get ahead Not enough patience Too frustrating Poor pay for hard work
Doctor	21	Sounds dull Can't stand suffering No stomach for it
Scientist	20	Too many atheists Boring Not smart enough Don't like science Don't like math
Garbage collector	17	Too dirty Poor pay
Secretary	15	Stuffy people Monotonous Hate typing See dull people Stuffy offices Don't like English
Laborer	14	Too hard Not good enough for me Insecure Too dirty
Factory worker	13	Monotonous Too hard Poor pay Cooped up
Gardner	13	Too dirty Rather do at home Don't like yard work Hate bugs Poor pay Hate pulling weeds Too hard
Nurse	12	Don't like to see suffering Can't stand shots Don't like blood

CHART 7 (Continued)

Occupation	Number	Reasons Rejected
Mathematician	10	Don't like math
Waitress	8	Too hard Poor pay
Salesman	8	Not forceful enough Uninteresting to do Have to meet people Poor pay
Farmer	8	Hated living on a farm Too little money Too hard Don't like animals
Artist	7	Not enough talent Not worthwhile Boring
Policeman	6	Too dangerous No future Like people too much
Bus driver	5	Can't stand buses Boring Crabby passengers
Mechanic	5	Not good enough for me Too dirty No talent
Janitor	5	Too dirty Boring

sion of the semester, they were also the most rejected occupations. Only one girl rejected the number one ‘‘glamour’’ occupation for teen-age girls, that of stewardess. No reason for the rejection was recorded.

Conclusions. Attitudes, whether engendered by adequate information, misinformation, or fancy, undoubtedly play a significant part in the career inclinations of ninth graders. With only an inadequate picture of the occupational opportunities of their community, they tend to judge occupations fancifully. In addition, with only a partially-formed self-concept they do not have a realistic picture of their own suitability for the occupations that attract them.

Work experience at the age of 14 is so limited that it has not as yet lowered the aspirations of the boys and girls. Eighty-four students had worked for pay outside their homes, with boys having more varied experience, encompassing a greater variety of jobs than the girls could claim. Jobs for both boys and girls were unskilled in nature. Few were related to the occupational aspirations of the students, although 26 students stated that they liked their jobs because they ‘‘ were learning a trade ’’ for their future. Actually, only four students later named career choices for themselves that were related to the jobs they had held.

Both boys and girls had more favorable things to say about the jobs they had held than unfavorable. Boys had more comments in both categories, including more write-in responses. A group of 27 students had no dislikes to record about their jobs at all, although every one of the 84 students holding jobs recorded some favorable comment.

‘‘Pay’’ was mentioned in favorable comments more often by boys than by girls, with girls choosing most frequently the ‘‘tasks’’ as an element they like about their jobs. Boys also named ‘ working with adults and being treated as an equal’’ more frequently as being important in liking a job than did the girls.

As for the things they disliked about their jobs, of paramount importance to the boys was ‘‘ always being told what to do.’’ The major objection of the girls was ‘‘the long hours ’’ they had to work. Also disliked by many girls was learning that other people were paid more for the same kind of work. Boys, more often than girls, objected to the work being ‘‘too hard.’’ Girls disliked the ‘‘long hours’’ and the children they had to babysit for: ‘‘ the kids were monsters.’’

The differences in the type of work they did probably contributed to the list of ‘‘likes’’ and ‘‘dislikes’’ about occupations and for the varying importance placed upon the items in the check list. However, there seems to be a maturation difference between boys and girls, with girls attracted to occupations or complaining about occupations more on an adult level than is true for the boys. There seems to be a need among the girls for enjoying the tasks and environment of the job that the boys do not share at this age to the same extent. Boys, on the other hand, as shown in this section of the survey and others, are attracted to occupations for financial reasons.

When asked to mark definitions of vocational ‘‘success,’’ both boys and girls seemed to pick the same items in almost the same numbers. In this category, boys along with girls chose ‘‘enjoying the tasks of the job’’ as a primary consideration. Not far behind in choice were ‘‘ earning the respect of one's fellow workers and employers.’’ The statement, ‘‘getting promoted on the job,’’ was third in preference for both boys and girls, but somewhat down the list from the two leading choices. Satisfaction on the job, shown in their choice of enjoying work tasks, and status in the eyes of others - very adult evaluations about what makes a job satisfactory - occurred to many.

The student responses for definitions of ‘‘success’’ were highly idealistic. A few boys and girls perhaps were facetious in responding to the item, ‘‘doing as little possible for the highest pay.’’ It is interesting to note that, throughout the study, about the same proportions of students chose items concerned with occupations that evoked envy in other people. In all, 18 students marked this response as a definition of ‘‘success.’’

Idealism seems to have been important in student responses to the question as to how people are promoted on their jobs. How much of this idealism comes from the practice of ‘‘ reward and punishment’’ children are familiar with at home and in school is hard to say. They overwhelmingly believed that people who win promotions do superior work, that they do more than is required of them, and that they have leadership qualities. There seemed to be little difference in responses between the sexes for all items except ‘‘they have leadership qualities’’ and ‘‘they are always quick to do more than is required of them.’’ In both cases, the girls seem to have been more optimistic than the boys. On the other hand, although the responses from boys and girls were few, the boys seemed to believe more than the girls that ‘‘ pull’’ might make a difference in promotional opportunity.

Apparent differences between attitudes of boys and girls were disclosed in the area of the survey concerned with the evaluation of what constituted a ‘‘good job.’’ Girls again seemed more interested in the environment of the job and the tasks to be performed than did the boys, as was noted in the student evaluations of the jobs they had actually held. Girls seemed less concerned with pay, again, than were the boys. Although again the ‘‘job that other people envy’’ was not important to the great majority of the students, to the boys this was a more important quality than it was to the girls. The original sample was reversed by the later follow-up study of the entire 9A class of Summer 1960, 445 students in all, with the [illegible] about routine work. In the original sample the boys outnumbered the girls in their statement that the "good" job would have " routine work that would not be too difficult." However, the second study disclosed more girls marking this factor than boys. In the original study those students who had chosen this aspect of a "good" job were, with only one exception, all in the lower ability or average ability classes. This information was not obtained for the second survey. In the larger second group, there was little difference between the boys and girls in their choice of "original, creative work," although the original study favored the girls in this response. The later study also did not bear out the strong trend in regard to an interest in working out-of-doors, with far more boys in the second study considering this essential to a "good job" than girls. In the first study almost as many girls named this as did boys.

When the students were asked to list five jobs apiece that they would not like and their reasons for the rejections, most students listed only two or three. A total of 70 different jobs was named. These fell roughly into two categories: those the students believed were low status and those requiring superior ability or training and which were the ones most considered by young people today. In other words, students rejected "blue collar" occupations which they considered " not good enough " for themselves and the high status professional occupations on grounds that " they hated" the school subjects connected with them. Some rejected occupations after an evaluation of their personal fitness, both academically and emotionally, but this was rare. The occupations which are most being urged upon students today and which are the ones students most picked for themselves when they listed occupations, were also the most rejected by others. The girls, thus, rejected secretarial work, nursing and teaching in the greatest numbers, although these occupations are invariably the most attractive to girls and are the highest status among the occupations traditionally considered by women. The occupation pupils were the most familiar with, teaching, was the most discarded from consideration. On the other hand, the professions of doctor and scientist were largely rejected by the boys, although these also attracted many boys. Occupations of "garbage collector," "laborer," and "factory worker" were low status in the eyes of the boys who rejected them. A few students rejected low status occupations on the basis of poor pay and insecurity, although there was no evidence that students had any real picture of the working conditions or rewards of such occupations. Girls did reject nursing, one of the traditional glamour occupations for women, although only one discarded "stewardess," the other important glamour attraction.

The study reveals some apparent differences in attitudes and beliefs, which may be true sex differences between boys and girls, or, more likely, culturally-inspired differences or maturation differences. Children by the age of 14 have learned the roles culture reserves for them and seldom consider occupations beyond the traditional ones. In addition, their considerations are extremely limited, with only 70 different occupations being mentioned. Girls seem more limited than boys, again likely as a result of cultural influence.

However, in their evaluations of success and promotional opportunity, there is no indication of any true sex differences. It might be possible that differences occur because of the differences in social strata from which the sample comes. This is a problem the survey did not answer.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 1, September 1962)

The Comparative Effects of Small-Group vs. Teacher-Led Discussion Sessions upon Student Achievement and Perception in Educational Psychology

WILLIAM J. GNAGEY
Butler University

EVER SINCE Levine and Butler (1952) pointed out that group discussion was more effective in changing behavior than the lecture method, there has been some doubt as to the necessity and efficiency of the teacher's participation in various classroom activities carried on in the name of learning.

If, as Hare (1952) concluded, there is an increase in participation in small groups as compared with large ones, one may question the efficiency of a teacher-led discussion involving more than 6 or 8 students at a time. If the teacher were to withdraw at intervals in favor of smaller divisions of his class, one might find his absence more of an aid to learning than his presence with the larger group.

On the other hand, some teachers and many students feel that unteachered small groups merely share their ignorance and waste valuable time.

These possibilities raise two major issues: 1) Is the teacher's presence ever indispensable for maximum achievement in such sessions? and 2) What does the teacher's presence do to the students' perception of the value of the course?

This study compares the achievement of two classes in educational psychology. One group was taught by a small group discussion method (SGD) and the other in a more conventional teacher-led class discussion (TLD). In addition, the investigation attempts to measure any differences in the students' perception of the practicality, interestingness and understandability of the course.

Some Definitions

The following operational definitions are proposed:

A teacher-led discussion (TLD) is carried out as follows: the instructor begins by inviting questions from the students about the chapter which has been assigned. Study sheets have previously been passed out containing essay questions and a list of terms considered important by the instructor.

When a student does ask a question, the instructor refers it to the whole class by saying, ‘‘Who can help us with this one?’’. If no other student volunteers an answer, the instructor answers the question himself, asking the questioner at the end whether he was satisfied with the answer. If not, the instructor pursues it further until the student claims understanding.

When a correct answer is volunteered, the instructor indicates his assent, but questions the asker as to whether he now understands. If not, the class is asked if anyone can explain further. If no further explanation is forthcoming, the instructor answers the question as before.

When an incorrect answer is volunteered, the instructor refers it to the class with the phrase, ‘‘Do you all agree?’’. Subsequent volunteer answers are handled as before until the questioner claims understanding of the point.

A small group discussion (SGD) is carried out as follows: the discussion is led by a duly elected chairman of each group of 9 or 10 students. With the help of the study sheets described earlier, they proceed more or less informally to help one another understand the information in the chapter under scrutiny. The instructor is available in the room for consultation with any group who needs his services, but is occupied with his own work in the interims.

Some Hypotheses

1. It is predicted that SGD classes should show greater achievement than TLD groups. The success of either of these discussion methods would appear to depend upon the student's willingness to get his questions before the group, and smaller groups are more conducive to such interaction. The teacher's availability to the SGD groups removes the possibility that they might be handicapped by text material too difficult for anyone of their number to comprehend sans an expert.

2. It is predicted that SGD classes should perceive the course to be more practical, interesting and understandable than TLD classes. With the greater participation possible in the smaller groups, it was felt that any semantic difficulties which might exist between the text and the students would be more readily worked through, thus making the author's ideas more precisely understandable to them.

The theory-practice gap was expected to be narrower in the small groups also since there would be greater freedom to compound practical applications than in a large class.

If it is true that student interest is increased by participation and involvement in the learning process (a current cliché), an SGD class should perceive the course as more interesting than the TLD group.

3. Whatever the outcome of the predictions, above, one would expect that whichever method was perceived to be more interesting, practical and understandable would be associated with the greater achievement. Otherwise some of our basic motivational assumptions need to be challenged.

Experimental Procedure

The experimental design included one independent variable: Method of Teaching (teacher-led discussion vs. small group discussion). The dependent variables were: 1)academic achievement and 2) student perception of the value of the course (practicality, interestingness, understandability).

The Subjects

Two college sophomore level Educational Psychology classes were selected for the experiment. Class A met at 1 p. m. daily except Tuesday and class B met at 2 p. m. in the same room on the same days. Table I shows the size and sex composition of the classes. They were assigned to these sections at random.

The Procedure

Groups A and B were both taught by the TLD method for 4 weeks. During this time, class work and assignments were kept parallel and they were each given two examinations. Use of the Chi-square median test showed that the groups did not differ significantly in achievement on these tests. The null hypothesis was accepted at the 80% level for test I, at the 25% for test II. In addition, a series of rating scales was administered at the end of the 4 weeks (see Figure 1). Analysis showed that the groups were not significantly different in their attitudes toward the course with respect to interestingness, practicality, or understandability. See Table II.

Beginning the middle of the 5th week, the method of teaching group B was changed to substitute the SGD method for the TLD method still used in group A.

Beginning the 6th week, the class work was carried on as indicated in Table III.

In class B, each group of 9 or 10 students was assigned so that it consisted of one A student, one D student, four B students, and three or four C students. The students were categorized as A, B, C or D on the basis of the results of the first two tests. The distribution of achievement ability in each group thus closely approximated that of the entire class.

Objective tests were administered again the 7th, 8th and 11th weeks to measure the level of achievement of both groups.

The rating scales measuring the 3 course value perceptions were administered the 8th week when the two groups were compared again.

Results and Discussion

No significant differences were found between the TLD and SGD groups with respect to achievement as measured by the examinations. See Table IV. From this it is concluded that this instructor (a variable which prevents us from generalizing to an unlimited universe), effected no appreciable facilitation of academic achievement and might as well have been correcting papers while the groups taught themselves. Though this is anything but reassuring to the college teacher, it should certainly challenge some preconceived notions about his function as a discussion leader of the type described.

Concerning the students' perception of the course's value, however, the picture is different. Results of the 8th week measurement of the student's evaluation of the course showed that the TLD group thought the course was significantly more practical than did the SGD group ($x^2 = 4.415$, $p < .05$). The TLG also showed a near-significant tendency to rate the course as more understandable than the SGD group

(x^2 = 3.825 and 3.841 = p. $<$ .05).

A much weaker tendency for the TLD group to rate the course as more interesting than the SGD group was non-significant (x^2 = 2.24, p. $<$ .10). See Table IV.

Again, the personality, delivery and general attitude of the experimental instructor would make us cautious about generalizing these findings broadly, but we were totally unprepared for results which forced us to conclude that student perception of the value of the course did not parallel achievement in the course.

Another possible implication is that, regardless of whether the presence of an instructor increases the achievement level of a group, he might be indispensable to morale, and thereby useful for that reason.

It must also be admitted that the SGD class may have learned some valuable behaviors that were not measured in this study. Such things as self-reliance, social skills, communication skills, etc., could very well tip the scales in favor of the SGD method if these are proper objectives for a course in Educational Psychology.

Perhaps the outstanding contribution that this study could make would be a vigorous stimulation of college teachers to test out some of their favorite procedures to see if they really are using their time most efficiently.

Summary

Two classes of college sophomore level Educational Psychology students were measured after exposure to two different teaching methods to ascertain whether a small group discussion method was superior to a teacher-led discussion method in producing achievement and positive value perception of the course.

While no significant differences were found between the groups' achievement, the TLD students thought the course was significantly more practical and tended to rate it more understandable than did the SGD group.

A weak tendency for the TLD group to rate the course as more interesting than did the SGD group was non-significant.

REFERENCES

1. Hare, A. "A Study of Interaction and Consensus in Different Sized Groups." *American Social Review*, 1952, 17, pp. 261-267.
2. Levine, J. and Butler, J. "Lecture Versus Group Discussion in Changing Behavior." *Journal of Applied Psychology*, 1952, 36, pp. 29-33.

TABLE I

SIZE AND SEX COMPOSITION OF THE TWO CLASSES COMPARED

Class	Number of Boys	Number of Girls	Total Number of Students
A	18	29	47
B	12	25	37

TABLE II

COMPARISONS OF EXPERIMENTAL AND CONTROL GROUPS BEFORE DIFFERENTIAL TREATMENT

Dependent Variable	Method	Group Scores* Below Median	Above Median	X^2	P.
Achievement					
Test I	TLD	22	25		
	SGD	15	21	.218	>.50
Test II	TLD	21	24		
	SGD	22	13	2.07	>.10
Practicality	TLD	12	26		
	SGD	9	21	.019	>.85
Interestingness	TLD	18	20		
	SGD	18	12	1.073	>.30
Understandability	TLD	10	28		
	SGD	12	18	1.434	>.20

*Scores exactly at the median are not used in the median test.

TABLE III

COMPARATIVE WEEKLY SCHEDULE OF CLASSES DURING EXPERIMENTAL PERIOD

Day of Week	Method Class A	Class B
Monday	Overview Lecture	Overview Lecture
Tuesday	-	-
Wednesday	TLD	SGD
Thursday	TLD	SGD
Friday	Test and TL Discussion of Results	Test and TL Discussion of Results

TABLE IV

COMPARISONS OF EXPERIMENTAL AND CONTROL GROUPS AFTER DIFFERENTIAL TREATMENT

Dependent Variable	Method	Group Scores* Below Median	Group Scores* Above Median	X^2	P.
Achievement					
Test III	TLD	22	24		
	SGD	20	15	.691	>.70
Test IV	TLD	22	23		
	SGD	18	14	.405	>.80
Practicality	TLD	17	25		
	SGD	22	12	4.42	.05
Interestingness	TLD	18	24		
	SGD	21	14	2.24	.10
Understandability	TLD	17	25		
	SGD	22	13	3.83	.06

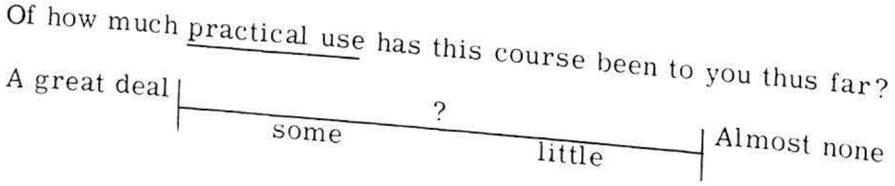
*Scores exactly at the median are not used in the median test.

FIGURE 1.

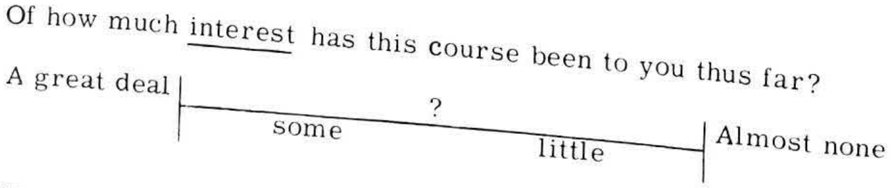
DEVICE USED TO MEASURE THE STUDENT'S PERCEPTION OF THE VALUE OF THE COURSE

Note: A line drawn vertically across the horizontal scale indicated each student's evaluation. Scores were the number of millimeters between the student's mark and the extreme right of the scale.

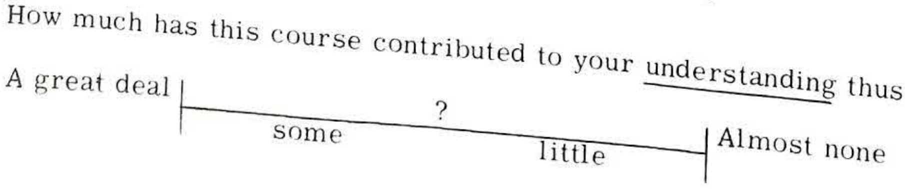
1. Of how much practical use has this course been to you thus far?



2. Of how much interest has this course been to you thus far?



3. How much has this course contributed to your understanding thus far?



THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 1, September 1962)

Do Parents Help Gifted Children Read?

DAN CAPPA and DELWYN G. SCHUBERT
Los Angeles State College

THERE has always been considerable discussion of what teachers should and shouldn't do to help children read. But it is a serious mistake to assume that reading is solely the responsibility of teachers and the school. A child spends more time at home than he does in school and a great deal depends on what happens there. What is the parents' attitude toward their children's reading? Do they feel the school is doing a good job of teaching reading to their children? Do they provide physical conditions and special reading materials for their children? Books? Magazines? Questions such as these are particularly important when considering the reading development of gifted children. Research shows that gifted children are mentally hungry and that their cerebral appetites often are met through wide reading. Since the school frequently is unable to provide enough free periods for gifted children to satisfy their insatiable desire to read, most of their free reading, if it is going to take place at all, must occur in the home.

With the cooperation of a number of experienced elementary teachers enrolled in the writers' classes, the authors were able to gather information about the home reading environment of 83 gifted intermediate grade children. The study grew out of questionnaires filled out by 29 fourth-, 33 fifth-, and 19 sixth-grade children. Two additional children were included, although no grade designation for them was available. Intelligence quotients ranged from 130 to 185 with an average of 137.8 Approximately one-third of the total number of children began to read before entering school.

Among other things, the children were asked how many books at home belonged to them, if magazines came to their home just for them, and the names of such magazines. They also were asked how much time they spent daily out of school, reading for fun.

When questioned as to how many books at home belonged to them, two of the 83 children answered with an ambiguous "many." One child failed to respond. The remaining 80, as evidenced in Table I, indicated having personal libraries ranging in size from 2 to 500 books. The average size was 41 books. Perhaps the most impressive aspect of these data is variability. Although no child reported owning no books, almost 20% of the children reported having small libraries (two to ten books); and a minority group indicated tremendously large libraries for this age level. In general, however, it is reassuring to find that 40% of the children own substantially large libraries of 20 to 100 books.

Forty-seven children in this study answered affirmatively the query, "Do any magazines come to your home just for you?" Table II shows that 21 different magazines were subscribed to by parents, with four: Children's Digest, Boys' Life, Jack and Jill, and Calling All Girls, enjoying a popularity accounting for 65% of the total number. As a sidelight, it is interesting to note that most of the magazines are juvenile in their appeal. A few adult magazines, such as Saturday Evening Post and Look, also are represented. In several instances, the magazine titles reported by the children are in error.

Table III shows the varying amounts of time the gifted children in this study devote to reading for fun outside of school. Seven children failed to commit themselves in this regard. But of the 76 who did respond, it is reassuring to find that well over half of them are devoting more than one hour daily to free reading outside of school.

"In the home, where do you do most of your reading?" To this question, the authors received multiple responses. Fifty-five children (56% of the total responses), answered: "my room." Another 27 children (28% of the total responses), said that they did their reading in the "living room." The remaining answers varied and included: "den," "kitchen," "spare room," "study," "in a corner," and "tree house." Three children failed to answer the question. The responses given, however, indicate that well over half of the gifted children in this study are provided with a suitable reading environment. Certainly it can be assumed that when a child reads "in his room," an optimal reading environment probably is to be found.

The last section of the study was devoted to studying parents' attitudes toward the reading habits of their children. Table IV shows that of a total of 217 responses, 176 involved criticisms of the children. On the other hand, statement "j" evoked 41 commendatory responses and statement "i" rated none at all. If we can put credence in the judgment of the children who filled out the questionnaires used in this investigation, the parents represented are happy with the school. None of them feels that the school is doing a poor job of teaching reading.

Statements "a" through "d" show that the parents are more interested in encouraging slow reading than fast reading. Of the two kinds of reading, oral and silent, the former is more subject to criticism. This is understandable, since shortcomings are more obvious when children read out loud.

Statement "e" shows that a substantial number of the parents (34%), want their children to do more reading. In a few instances, according to "f," parents have juvenile bookworms who constitute a problem by doing too much reading.

Not only are these parents interested in having their children read more, but, according to statement "g," 22% of them would like to see more difficult books read. Only one parent in the entire sample felt that his child should read easier books. Apropos, too, is the parents' desire to have children include in their reading a greater variety of books (statement "p").

Statements "k" and "m" don't prove to be of concern to parents but "l," on the other hand, reflects the current interest in phonics. Eighteen percent of the parents feel that their children "should do a better job of sounding out words." This tends to corroborate an earlier study by one of the authors, showing that gifted children do not make much use of structural or phonetic analysis in attacking unknown words.

About 28% of the parents are concerned with the inability of their children to remember what they read. Their preoccupation with statement "n" may be a reflection of the layman's conception that an educated mind is encyclopedic in nature. This insistence on retention has been given impetus by the numerous "Champagne for Caesar" types of television shows which glorify the ability to retain information.

The last item, "q," involves nine miscellaneous criticisms of children's reading such as: "you shouldn't read at the dinner table," "you should appreciate poetry more," "you should get more interested in books," etc.

In summary, it seems evident that the parents of the gifted children in this study are assuming responsibility in extending the reading skills of their children. The majority of them are buying books and subscribing to magazines for their children to read. They are providing time and a place in their homes where free reading can take place. As a group, parents are critical of their children's reading and are interested in raising the quantity and quality of the reading done by them. Parents do not, however, level any criticisms against the school.

TABLE I

HOW MANY BOOKS AT HOME BELONG TO YOU?

Number of Books	2-10	12-20	21-40	50-101	105-500
Children Responding	20	21	16	18	5
Average per Child	41 books				

TABLE II

WHAT MAGAZINES COME TO YOUR HOME JUST FOR YOU?

Name of Magazine	Frequency
Children's Digest	17
Boys' Life	8
Jack and Jill	5
Calling All Girls	5
Movie Magazine	2
Children's Friend	2
Child's Life	1
Around the World	1
Reader's Digest	1
Linn's Weekly Stamp News	1
Look	1
Girls' Life	1
Comic Books	1
American	1
Children's Activities	1
Photoplay	1
Sport	1
Saturday Evening Post	1
Wee Wisdom	1
Junior Natural History	1
American Girl	1

TABLE III

ABOUT HOW MUCH TIME DO YOU SPEND DAILY, OUT OF SCHOOL, READING FOR FUN?

Length of Time in Hours	1/4 - 1	1 - 2	2 - 3	3 - 5
Number of Children	30	29	15	2
Average per Child	1.16 hours (70 minutes)			

TABLE IV

CHECK ONE OR MORE OF THE FOLLOWING STATEMENTS THAT BEST DESCRIBE ATTITUDES OF YOUR PARENTS TOWARD YOUR READING HABITS

a.	6	you should read faster orally
b.	17	you should read slower orally
c.	7	you should read faster silently
d.	9	you should read slower silently
e.	28	you should read more
f.	5	you should read less
g.	18	you should read harder books
h.	1	you should read easier books
i.	0	you are not being well taught to read at school
j.	41	you are being well taught to read at school
k.	6	you should know more words when you read
l.	15	you should do a better job of sounding out words
m.	2	you should understand better what you read
n.	23	you should remember more of what you read
o.	13	you shouldn't read in bed
p.	17	you should read a greater variety of books
q.	9	others: ______

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 1, September 1962)

Scholastic Aptitudes, Vocational Interests, and Personality Characteristics of Journalism Students

EDMOND MARKS, JOHN D. VAIRO and MARTIN L. ZEIGLER
The Pennsylvania State University

JOURNALISM educators and the field of Journalism in general are becoming increasingly concerned about the fluid state of journalism enrollment, as reported by schools of journalism throughout the country. Although total enrollment remains fairly constant, the withdrawal from, and transfer to journalism is quite high, according to trends reported over the past several years.

Another disconcerting trend to journalism educators is the growing predominance of women in relation to men who are entering schools of journalism. Practical considerations for women, such as marriage and motherhood, indicate a possible instability in the status of journalism as a profession.

Realizing the far-reaching implications of these trends, the School of Journalism of the Pennsylvania State University is gathering data to be used both in the evaluation of the various hypotheses, and to indicate new hypotheses regarding these trends. Beside these immediate enrollment trends, the project seeks to establish the concept or image of journalism which high school students and entering college freshmen have of such a career.

The study reported here was conducted by Student Affairs Research in collaboration with the School of Journalism as part of that school's research project. The present study attempts to determine if there are any significant differences in scholastic aptitudes, interest patterns, or personality characteristics between those students enrolled in journalism and those students who had enrolled but later withdrew from that school.

Subjects and Procedure

Scores on aptitude, interest, and personality tests were compiled and evaluated for the 339 individuals comprising this study. Of these 339 students, 160 were male and 179 were female. The students were classified into two groups as to: 1) those remaining in the School of Journalism (N = 187), and 2) those students withdrawing from the School of Journalism (N = 152).

For those students remaining in the School of Journalism, the data revert to the spring semester of 1955, being compiled at semester intervals until the fall semester of 1958, thus giving eight semesters of data; i.e., those individuals entering the School of Journalism as freshmen in the spring of 1955 and remaining therein comprise the eighth semester scores, with the following semester students comprising the seventh scores, with students entering the School of Journalism in the fall semester of 1958 being regarded as first semester students. This method assured that no single subject appeared in both the "remaining group" and "withdrawal group."

Those individuals comprising the withdrawal group were categorized according to the semester when they withdrew from the School of Journalism.

Aptitudes of incoming freshmen were measured prior to any formal course enrollment by the Pennsylvania State University Academic Aptitude Examination (1). This test is composed of four sub-tests labeled Vocabulary, Paragraph Reading, Arithmetic, and Algebra. These four sub-tests are then summed to give a four-part total score. Because the number of students comprising the second, fourth, and sixth semesters in the remaining group, and fifth, sixth, and seventh semesters in the withdrawal group, was so small as to make analysis and comparison of the individual sub-group scores meaningless, the groups were divided in half on the basis of first through fourth semester and fifth through eighth semester standing. Inter-group and intra-group comparisons were then made to determine the existence, if any, of distinguishing differences in scholastic aptitude.

Interests of the students were determined by the Strong Vocational Interest Blank, which consists of two separate inventories—one for males and one for females. The male inventory is scored for 36 different occupational scales, whereas the female inventory is scored for only 25 occupations. The greater portion of these occupational scales, however, are not found on both inventories, so that direct comparison between male and female interests is limited. For a more detailed description of the Strong Vocational Interest Blanks, the reader is referred to the manuals prepared by Strong (2, 3).

Since the nature of journalism necessarily involves literary interests, this group of scales was of particular interest to the investigators. The Literary group includes the vocations: Advertising, Author-Journalism, and Lawyer. Author-Journalism was not included in the earlier usage of the Strong Vocational Interest Blank, therefore no intra-group comparison could be made.

In an effort to obtain finer discrimination and facilitate inter-group and intra-group comparisons, both initial groups were divided into two additional groups on the basis of 1) first through third semester, and 2) fourth through eighth semester groupings. This division was effected on the premise that the greater portion of withdrawals will occur in the first three semesters. Since the initial group had been previously divided into male and female, this third division yielded eight groups for comparison of interest patterns.

Personality traits were assessed by means of the Bernreuter Personality Inventory (4). The personality traits as reported in the results are stability, self-sufficiency, and dominance.

Results and Discussion

Aptitudes: Scores of the Penn State Academic Aptitude Test are reported in terms of mean four-part total scores obtained by the various semester groups. The results are reported in Table I (those students remaining in journalism), and Table II (those students withdrawing). The number (N) comprising each group is given beneath the stated semester.

From an inspection of Tables I and II, a difference in mean aptitude scores is discernible between those students remaining in the School of Journalism and those students withdrawing from journalism. The mean aptitude score of the remaining group is ten points higher than that of the withdrawal group. As such, scholastic ability in relation to course requirements, may be a factor in a student's decision to withdraw. It should be noted that although the difference is fairly large, prediction of withdrawal based entirely on aptitude scores alone, without reference to other cogent factors influencing such a decision, would be ill-advised.

Differences within the two groups were so small as to be assumed to have little significance.

Interests: A problem encountered in this study in analyzing and evaluating the scores yielded by the Strong Vocational Interest Blank, was the fact that over the period of time that this study encompasses, the scoring procedure employed by Student Affairs Research changed, so that more vocational scales were added. This negated direct comparison of occupations except for those scales found on both the older and newer usages of the Inventory.

A second problem involves the fact that not all the students comprising this study had taken the Strong Vocational Interest Blank, again affecting any direct comparison.

The results are reported in terms of the percentage of each group achieving the standard established by Strong, i.e., scores of A or B+. Because of the difference in content between male and female forms of the Inventory, separate results are reported.

All the male groups scored highest in the Literary and Sales areas,followed closely by the Business and Social Service categories, in that order. Those categories having the lowest percentage of corresponding scores were the General Professional and Science-Engineering categories. Advertising achieved the highest score for all male groups (averaging 70%), except for the fourth-eighth semester men who remained in journalism, where it dropped to 44 percent. This is worth noting, since Advertising is one of the two departments of the School of Journalism, with 56 percent of those males remaining in journalism being enrolled in this department. Unfortunately, no score is available for the Author-Journalism category for this group.

The women's scores show no single concentration of interests, although the Literary area, as a group, showed a higher percentage of A or B+ scores than any other group, but this too was considerably lower than the male scores, averaging only 25 percent.

From an analysis of the data, a characteristic interest pattern of all students entering the School of Journalism has been established. The highest scores distribute themselves primarily among the Literary and Sales areas, and secondarily among the Social Service and Business groups. Both the students remaining in and those withdrawing from the School of Journalism exhibited the least interest in General Professional and Science-Engineering occupations. Although Literary scores are generally the highest, the high percentage of scores in these other areas indicates considerable diffusion of interests throughout the non-science areas.

Although a characteristic interest pattern of students entering journalism is evident, prediction as to whether a student will remain in or withdraw from that curriculum based on interest scores alone cannot be made. The differences in interests of the various groups is so slight as to make any discrimination between them on the basis of characteristic interests unwarranted.

Personality Traits: The results of personality trait assessment are reported in terms of the percentage of each group falling into the trait classification established by Bernreuter.

The results indicate that men in the fourth-eighth semester group who remained in journalism appear slightly more emotionally stable and slightly more dominant, but slightly less self-sufficient than any of the other male groups. Males, both fourth-eighth semester men in and out of journalism, had a higher percentage of stable scores than did either of the first-third semester groups; however, caution is prudent in interpreting these results because of the small number of cases (N = 11) comprising the

TABLE I

MEAN FOUR-PART TOTAL SCORES OF STUDENTS REMAINING IN JOURNALISM ON THE PENN STATE ACADEMIC APTITUDE TEST

	1-4 Semesters (N = 102)	5-8 Semesters (N = 85)
Mean	138.6	136.8
S. D.	26.8	26.7

TABLE II

MEAN FOUR-PART TOTAL SCORES OF STUDENTS WITHDRAWING FROM JOURNALISM ON THE PENN STATE ACADEMIC APTITUDE TEST

	2-4 Semesters (N = 133)	5-7 Semesters (N = 19)
Mean	127.8	124.3
S. D.	26.5	23.1

fourth-eighth semester group out of journalism.

The women's scores exhibit an almost opposite pattern to the men's scores. Fourth-eighth semester women who remained in journalism had a higher percentage of scores falling into the self-sufficient category than any of the other groups but scored higher in submissiveness than either the first-third semester women in journalism or fourth-eighth semester women out of journalism.

In comparing the male and female results, certain trends are discernible. In this study, the women entering the School of Journalism scored higher in dominance and lower in dependence than did the men entering journalism.

Again, the slight differences in personality characteristics of the various groups, as measured by the Bernreuter Personality Inventory, preclude any prediction of withdrawal tendency based on the results of personality assessment.

Summary

The results of aptitude, interest, and personality traits tests were accumulated for 339 students who had enrolled in the School of Journalism since the spring semester 1955, ending with the fall semester of 1958, in an effort to determine characteristics peculiar to this group. In addition, the group was also divided according to whether the students remained in, or withdrew from this curriculum, and also male and female, in an effort to derive distinguishing characteristics, if any.

From the results of aptitude testing, we can draw no conclusions regarding any general trend within groups. Although the withdrawal group scored consistently lower than the remaining groups, other selective factors are probably involved in decisions regarding withdrawal.

An analysis of interest scores indicates that students entering journalism are primarily interested in literary and sales vocations and secondarily in social service and business occupations. The least corresponding interest scores were related to general professional and science-engineering vocations.

As to personality traits, males and females in the fourth-eighth semester category remaining in journalism exhibit almost opposite patterns. The males score higher in dominance and emotional stability, but the women score higher in self-sufficiency. This distinction is also true for the fourth-eighth semester male group remaining in journalism as compared to all other male groups.

Differences in interests and personality traits of the various groups are so slight, however, as to make discriminations between them on the basis of measured interest patterns and personality traits unwarranted.

REFERENCES

1. Moore, B. F., and Castore, G. F., The Pennsylvania State University Academic Aptitude Examination, The Pennsylvania State University, 1947.
2. Strong, E. K., Manual for Vocational Interest Blank for Men, Stanford, California, Stanford University Press, 1951.
3. Strong, E. K., Manual for Vocational Interest Blank for Women, Revised Blank (Form W) and Scales, Stanford, California, Stanford University Press, 1951.
4. Bernreuter, R. G., Manual for the Personality Inventory, Stanford, California, Stanford University Press, 1935.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 1, September 1962)

Case Studies in the Changing Roles and Statuses of College Teachers

JAMES A. MOSS
Union College

THIS ARTICLE developed out of an intensive study of the Negro college teacher in predominantly white colleges in New York State.[1]* This newly emerging segment of the academic profession affords an interesting contrast with colleagues in other teaching arrangements. Presented below are profile comparisons of four major groups of college teachers in the United States: (1) the white teacher in the predominantly white college,[2] (2) the Negro teacher in the Negro college,[3] (3) the white teacher in the Negro college,[4] and (4) the Negro teacher in the predominantly white college.[5] Since three of the studies used as the background for this article were carried out without reference to the author's own study, the data presented are not uniformly weighted. Where these studies are comparable, information is provided in the following areas: (a) the social origins of college teachers; (b) motivation for entering the teaching profession; (c) the training, recruitment, and position of college teachers in the social structure, and (d) the self-perceptions of college teachers as well as the images of them held by others in the college structure.

The White Teacher in the White College

In the predominantly white college, the white teacher most frequently is male, holds the rank of assistant or associate professor, is from the lower middle class status, born in one of the eastern or north-central states, thirty-seven years of age, and is married with one or two children. He received his M.A. and/or Ph.D. degree from one of the major universities; has taught at only one other college, and has been in his present teaching position from six to ten years. With varying degrees of identification with the profession, he enters as one of the many confused, timid and introverted students afraid to leave the relative security of undergraduate life and who found the three or four added years of Ph.D. work to be the line of least resistance. Some find "in teaching a meal ticket to the library cubicle"[6] or the "price of the key to the laboratory";[7] still others seek the social status accorded to members of the college teaching profession. Few are described eager to teach and committed to making teaching the central goal in their lives. Contrary to some popular notions that he is mediocre and dull, the white teacher, in all probability, is drawn from those student ranks containing the more superior minds; and his entry into the college teaching profession is a personal choice coupled with the fact that his position carries with it high social prestige and value.

In his rise in the academic hierarchy, he tends to develop more of the "intellectual" than the "social" graces; and although his membership in the college teaching profession serves to advance him into the upper ranks of the middle class, both in income and in social rank, he still falls into a bracket lower than that of successful doctors, lawyers and businessmen whose "style of life" (mode of living), his most nearly approximates. While not highly class conscious, he is envious of the performance of his colleagues, but is relatively free from bigotry, bias, careerism, and other factors injurious to the disinterestedness necessary for the best professional performance.[8]

Since only one in every fifty generally goes on for the Ph.D. degree, his formal graduate training terminates with the M. A. degree. However, if he has the Ph.D. degree, it was earned at an institution other than the one which conferred his bachelor's degree and he has shown preference for one of thirty leading graduate schools. If his position is within a private or church-related college, he is also a graduate of that institution and was given preference over someone from outside his Alma Mater. In most cases, he received appointment through friends, relatives, and/or former teachers in his employing institution. Once a member of the faculty, however, he is alerted to the roles and attitudes expected of him if he aspires "to move up the ladder" and, as a teacher, surmises early in his career that retention and promotion depend less upon good teaching than upon the prevailing character of the department, the idiosyncracies and personal preferences of its chairman, the ability to "make himself liked" by other colleagues, and the size and nature of the institution in which he performs.

---

* Footnotes will be found at the end of the article.

Notwithstanding, selection and appointment is generally based upon qualities such as: "ability and training in teaching and research, intellectual calibre, capacity for leadership, general reputation, originality, personality and character, and probably future performance."[9] Whether recruitment is from without or from within the institution, the usual beginning position, however, is on the lowest rung of the academic ladder - that of instructor. Promise of tenure, or the lack of same tends to remind him that is employment is insecure and that conformity is the surest path to job security. Although as an instructor tenure is often short-lived, there is much anxiety and uncertainty; turnover is high, and pay is low. "Short-term" appointments prolong his feelings of job insecurity.

Some satisfactions accrue to him through his affiliation with the superior staff found in a major university or college. Also, this satisfaction may be enhanced by greater opportunities for democratic participation in decision making, favorable working conditions under an academically minded president, and enjoyment of adequate provisions for tenure. Conversely, however, his dissatisfaction may be greater in the lesser institution where the administration tends to be more authoritarian; where tenure provisions are lacking; where personnel turnover is relatively high, or staff competence low; and where the president reflects more interest in "bricks" than in "brains." Moreover, dissatisfaction may stem from incompatibility with other colleagues; inability to work harmoniously with his departmental chairman; depression arising from low rank and low salary; or the imposition of a heavy teaching load draining energies or curtailing time otherwise available to him for additional study, research and publication.

By the time that he becomes thirty-eight, he has achieved the rank of associate professor and is accepted as a member of the group both by his colleagues and by the community in which he lives. As a forty-year-old, he may well have attained the rank of full professor. At such a time, he may feel himself well beyond the stigma of non-promotion and may well anticipate some easing of his labors, through university-provided research and clerical assistance. In his role as a departmental head, appointed by the president, elected by departmental members, or rotated among staff members above a certain rank, he exerts decisive influence upon appointments, promotions, and assignment of courses and teaching loads. In this capacity his greatest stresses and strains from within the college are likely to revolve around departmental rivalry for students, funds, equipment and expansions and departmental tendencies toward isolation from the larger college. While he is resentful of power exercised by superiors, he relishes and seeks authority in much the same way as other parts of the social structure. His status, as perceived by him both from within the institution and from the community, is a reflection of his income level and, from this point of view, it appears that no way exists for him, outside of private and governmental research funds, to completely overcome income as a social denominator in a society where wealth is such an important criterion for status.

The Negro Teacher in the Negro College

In the Negro college, the teacher most likely is male, was born in the South, and educated in a racially segregated high school and college. He comes from a low-income family and has worked his way through college. Even if he is included in the twenty-four percent who hold the Ph.D. degree, he does not perceive his social origins and background as deterrents to his moving up the ranks, and succeeds in bridging the gap between the bachelor's degree and the Ph.D. degree in approximately twelve years. Having completed advanced work for the Ph.D. degree in a white institution, he expresses satisfaction with his choice of an academic career. However, he finds himself in a favored position only if he has received the doctorate degree, otherwise he is carrying a course load of more than twelve hours, is without tenure, and is considering changing his position to improve his salary or status. The higher he climbs, particularly if he holds the Ph.D. degree, the more mobile he becomes, in the knowledge that the shortage of Negroes with this degree makes his bargaining power exceptionally high.

Nevertheless, irrespective of rank as a new recruit or as a senior member of the faculty, he is left alone to learn his duties and responsibilities and is seldom pressured to undertake research. Although he may earn more than do women of equal rank, he earns less than if he were in a private non-denominational college. With the one exception of being employed in a state institution, his highest salary accompanies his rank as associate professor. Consequently, as he looks up the ladder of success, he feels that he does have the respect or recognition from either the president or his immediate superiors and that if his success is measured in terms of salary, he is noticeably unsuccessful. Resigned to his present plight, he is convinced of being at a "dead end" with very little hope of ever becoming a professor at a major university.

The White Teacher in the Negro College

In the main, the white academic man in the Negro college is either foreign-born or born in or near New York. Chances are two out of three that he holds the doctorate degree and is about forty-three years of age. He is teaching English, religion, and foreign languages and has been teaching in a Negro college for about two and one-half years. He has no special motivation for teaching in a Negro college, and expresses the opinion that "color" as such has no particular meaning for him His attitude towards integration is that anyone should be allowed to go to any college; integration from both directions is necessary, and that as integration progresses, Negro colleges *per se* will disappear and those remaining will be completely integrated. His cognizance of the fact that his students come from poor economic and/or academic backgrounds seems to increase his capacity for patience and understanding and fosters "particularly close and rewarding" relationships among them.

Although his relations with students, administrators, and other colleagues are good, only a few of his relatives and friends view his teaching arrangement favorably. In spite of the fact that his living in a Negro community offers "frequent and enjoyable interracial social gatherings," his prestige is still somewhat lower than the white teacher teaching in the predominantly white college, with the result that he is not totally integrated either in the local white or Negro community.

The Negro Teacher in the White College

In the white college, the Negro teacher may be male or female; born in the South, although in increasing numbers born in the North of southern parents who have migrated in search of improved social and economic opportunities; is in his mid-forties, married; has one or two children; and has undertaken his graduate studies either while single or while married with no children. His mother at present teaches school or has taught in the past; his father, slightly less educated than his mother, holds a relatively secure white-collar position. In spite of the fact that parents and former teachers have influenced his selection of college teaching as a vocation, his own desire is consistent; and his strongest voiced motivation for entering the profession is the "kind of mind" found on the college level. Having taught in a Negro college where tenure provisions, salary, promotion, lack of concern for scholarship, and authoritarianism of college administrations were not favorable, he is now concentrated, for the most part, in the large, publicly-supported college and expresses a greater desire to "go up the ranks" within his present institution than to shift into either another college or another profession.

Although he verbally denies that only a few white college teaching posts are open to him, and that certain concessions are made by him in both salary and rank in order to secure his appointments, he entered into his present position with a Ph.D. degree from a major white institution, with both the master's and bachelor's degrees from outstanding Negro institutions, and with at least ten years of teaching experience in reputable white and Negro colleges. His considerable experience is accounted for largely by the fact that he entered the profession between the ages of twenty and twenty-five, an age somewhat younger than most beginning college teachers.

Although he considers himself fully adequate to the demands of his assignment, he either underestimates his relative position with his professional peers or approximates closely the assessment accorded him by his department chairman. In a minority of instances, his self-evaluation is higher than that given him by his immediate superior. At times he feels he is "being put through the test" by his white colleagues; however, his major grievances revolve around the issues of salary and promotion. These difficulties, often more real than fancied, are interpreted as being associated with his minority status; and, on occasion, his anxieties are buttressed by evidence from his administrators who confirm his longer than usual tenure in rank and the Ph.D. as the basic requirement for his initial appointment.

Despite his feelings of full acceptance as a person by his colleagues and students, and that in at least a few instances his being a Negro may operate in his favor with respect to promotions, he also believes that in order to go up the rank he needs not only more publications but also more support from and recognition by the "right" people in the social structure. Reflecting that the pace of integration both in the hiring of Negro teachers at the college level and in the acceptance of Negroes into the larger white society has been proceeding too slowly, he voices very strongly the opinion that the responsibilities for bringing about change in these two situations rests not only upon him, but also upon other citizens or professional educators.

Summary

There appears to be a similarity in the parental social status of the white teacher in the white college, the Negro teacher in the white college and the white teacher in the Negro college. All three of these groups identify their family socio-economic background as lower middle and middle class. Only the Negro teacher in the Negro college is reported to derive from families which occupy the lower income brackets.

Motivation for teaching appears consistently high for both Negro teachers in white colleges and white teachers in Negro colleges. No such uniformity of feeling about the teaching profession seems present among white teachers in white colleges, and the Negro teacher teacher in the Negro college appears as the least motivated of the four groups.

High motivation for teaching seems related also to the number of teachers who continue their professional training through the Ph.D. degree. Two out of three Negro teachers in white colleges and white teachers in Negro colleges hold this advanced degree. Among Negro teachers in Negro colleges, one out of four holds the doctoral degree, while those who have attained this level of education among white teachers in white colleges numbers one out of fifty.

Upon achieving permanent status, all but Negro teachers in Negro colleges appear satisfied in their assignment. In this latter group, mobility among the holders of the Ph.D. degree is especially high.

Full participation in the social structure of the college seems chiefly confined to white teachers in white colleges, to a lesser degree to Negro teachers in Negro colleges and to a minimal degree to both white teachers in Negro colleges and Negro teachers in predominantly white colleges. While neither Negro teachers in white colleges nor white teachers in Negro colleges are represented to any extent in the administrative ranks of their institutions, there is no indication that a "ceiling" exists with regard to their opportunities for professional advancement.

There appears to be a striking resemblance between the Negro teacher in the white college and the white teacher in the Negro college. In both instances they are minority members in institutions controlled by a racial or ethnic group other than their own. Both groups consist of individuals who feel secure enough to break away from teaching traditions in their own group and test their abilities in unfamiliar, new teaching environs. They come from as comfortable socio-economic backgrounds as their white colleagues in white colleges. They are likely to have come into their present teaching assignment with previous experience and with somewhat better preparation than their colleagues. With the exception of their fellow teachers in Negro colleges, they are as satisfied in their present positions as other teachers who have achieved permanent status. Neither of these two groups is significantly represented in the administration of their colleges, but there is no sentiment expressed to the effect that they are excluded from holding such posts.

Finally, both groups express strong concern for more rapid integration of teachers at the college level.

FOOTNOTES

1. Articles previously published based upon this research include: "The Integration of Negro Teachers in Predominantly White Colleges," *Journal of Negro Education*, Howard University, Washington, D.C., Fall, 1958; "The Utilization of Negro Teachers in New York State," *Phylon*, Atlanta University, Georgia, March, 1960; and "Currents of Change in American Race Relations," *British Journal of Sociology*, September, 1960.
2. Wilson, Logan. *The Academic Man*. New York, 1942.
3. Thompson, Daniel. "The Teacher in the Negro College: A Sociological Analysis," unpublished Ph.D. dissertation, Department of Sociology, Columbia University, 1956.
4. Decker, Paul. "A Study of 'White' Teachers in 'Negro' Colleges, *Journal of Negro Education*, Vol. 4, No. 4, Fall, 1955, pp. 501-5.
5. Moss, James Allen. "The Utilization of Negro Teachers in the Colleges of New York State," unpublished Ph.D. dissertation, Department of Sociology, Columbia University, New York, 1957.
6. Wilson, *op. cit.*, p. 7.
7. *Ibid.*, p. 7.
8. *Ibid.*, p. 115.
9. *Ibid.*, p. 118.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 1, September 1962)

The Effect of Musical Distraction upon the Performance Efficiency of Children

ALVIN ENGEL and CYNTHIA ENGEL
Roslyn Heights, New York

SEVERAL STUDIES have been reported on the effects of music upon the efficiency of people within their specific areas of function. Some of this work has been done in industry.

Among these is the study by Kaplan and Nettel (3) in cooperation with the British Broadcasting Company and other researchers. Their results showed that music can aid production by boosting tired workers, acting as a mental tonic, relieving boredom, minimizing conversation and cutting down on absenteeism. The only qualification mentioned was that the music was to be properly chosen.

Wormer (4), in his work with the Muzak Corporation in the United States, found that among workers, the music that was rated the highest was waltz music and popular dance music. The Muzak Corporation made a study of 19,546 workers. Only a little less than four percent of this number had any objection to having music at work.

In 1937, Frederick (1) studied the influence of musical distraction on reading efficiency. His study was done with college students in several periods of distraction. His results showed that the non-distracted group was more efficient, and that the musical distractions affected most of the students of higher intellectual levels.

One of the more recent studies in this area was done by Freeburne and Fleisher (2) at the Ohio State University. They used five groups of forty students each for their study. Four of the experimental groups read with music—jazz, classical, semi-classical and popular. The one control group read without music. Their results showed that the jazz group read more quickly than the other groups. However, the scores on reading comprehension tests showed no difference among all five groups.

The present study differs from the aforementioned studies in the following manner: first, these studies were done with either college students or adults. Second, the task to be measured was the efficiency and comprehension of reading material. Third, the distraction employed—jazz and popular music—might have a tendency to involve subjects emotionally, and to elicit daydreams from the students of college age. This unmeasured variable may have as much effect upon the distraction of attention as the experimental music itself, thus reducing efficiency even more.

Problem and Hypotheses

The present study investigated the effect music has upon the efficiency of performance of fifth grade students doing routine arithmetic combinations.

The hypothesis to be tested is: There is no difference in performance efficiency of a group of fifth grade students doing simple arithmetic problems under the conditions of musical distraction and silence.

Method

The tests used to measure arithmetic learning consisted of twenty number problems involving the four processes—addition, subtraction, multiplication, and division. There were five problems for each process. No written material was used in any of the problems. They all appeared in their algorism form on dittoed sheets.

The tests were given every Wednesday morning for eight consecutive weeks, before boredom, restlessness and hunger set in. They were alternated one week with musical background and one week with silence. The duration of each test was forty-five minutes. In order to insure uniformity of time allotted to each test, no paper was collected before the testing time was up. All material was taken from the fifth grade arithmetic textbook currently in use.

The musical distraction came from a phonograph placed in the rear of the classroom. The selection played was the Piano Concerto Number 2 by Chopin.

In order to keep the tests as objective as possible, the material chosen was such that it could be judged right or wrong. All ambiguity such as is inherent in reading for comprehension was therefore eliminated. All problems were in numbered form instead of worded, thus eliminating the possibility of discrimination against the poor reader.

The twenty-five subjects participating in this study were from a regular fifth grade class chosen at random. All participants were working on a fifth grade level in arithmetic. They ranged in age from ten to eleven and a half years.

TABLE I

ANALYSIS OF VARIANCE

Source	Sum of Squares	D. F.	M. S.	F.	Null Hypothesis
Between Methods	54.08	1	54.08	<1.00	Accepted
Within	3378.24	48	70.38		
Total	3432.32	49			

TABLE II

ANALYSIS OF VARIANCE - COMPLEX FACTORIAL DESIGN

Source	Sum of Squares	D. F.	M. S.	F.	Null Hypothesis
Methods	13.52	1	13.52	3.26	Accepted
Pupils	760.33	24	31.68		
Trials	23.42	3	7.81		Accepted
M × P	84.23	24	3.51		Accepted
M × T	23.36	3	7.79		Accepted
P × T	181.83	72	2.53		Accepted
Residual	298.89	72	4.15		
Total	1385.58	199			

Results

An analysis of variance was computed for the above data, and a further analysis was done for a complex factorial design including tabulating the sum of squares between trials, pupils, conditions and their respective interactions. None of the results had any statistical significance. The tabulations may be seen in Tables I and II.

This indicates that there is no difference between the experimental condition of music and the control condition of silence. It would seem, then, that the original hypothesis as stated, namely, that there would be no appreciable difference between the efficiency of fifth grade students doing arithmetic combinations under conditions of silence and musical distraction, is accepted.

Discussion

While it is accepted that music makes working conditions pleasant, it is not really that factor that is the central issue here. Since it seems that the presence of music does not decrease the efficiency of an average fifth grader to do number problems, the main objective is not to make their school day more pleasant, although this has its positive values. There is a more crucial issue here, and that is the exposure of these children to "good music." This can be considered a planned and concentrated effort to expose these children to culture without announcing it as such.

It has been fairly well established in child psychology that children like and prefer that which is familiar. It is true that only those who are especially interested in music will ask the name of what they hear. They are all hearing it, however, even though they may not be attending to it. Through repeated hearings such as this, it becomes part of their total gestalt of what they know and recognize. Recognition gives it a positive tone, and removes it from the sphere of the "adult, long-haired stuff for squares." These experimenters feel that such a positive attitude toward classical music may be carried over into adolescence and later into adulthood.

In this connection, in some unpublished work done with kindergarten children, it was found that they would hum themes from Beethoven symphonies while building with blocks, as easily as they would the usual primary songs, after they had had repeated exposure to the classics in a simplified form. They could play their rhythm band instruments to Katchaturian themes as easily as they could to primary grade marches.

Perhaps young children are much more capable of learning to appreciate the fine arts through incidental learning than we usually are aware. With all of the recent emphasis on the sciences, it is important for educators and educational consultants to make certain that the arts are not lost in the shuffle.

Summary

In this study, 25 fifth-grade students had four trials each, under conditions of musical distraction and silence, to do 20 arithmetic combinations. Since the same group of 25 children were tested under both conditions, there was no problem of matching groups. Analysis of variance was computed for the differences between conditions, pupils, trials and their respective interactions.

None of the resulting F ratios were statistically significant. Therefore the null hypothesis is accepted, namely, that there is no real difference in the ability of fifth graders to do arithmetic combinations under conditions of music and silence. The meaningful application of these findings to the education of our children is discussed.

REFERENCES

1. Frederick, P., "The Influence of Musical Distraction upon Reading Efficiency," Journal of Educational Research, 1937, 31, 264-271.
2. Freebourne, C., and Fleischer, M., "The Effects of Musical Disturbance upon Reading Rate and Comprehension," Journal of Educational Psychology, 1952, 43, 101-109.
3. Kaplan, L., and Nettel, R., "Music in Industry," Biology of Human Affairs, London, 1948, 13, 129-135.
4. Wormer, E., Work by Muzak, 1945-47, A research study of its effectiveness and acceptance, N. Y. Muzak Corp. 1948.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 1, September 1962)

Parent-Teacher Relationships in the Early Nineteenth Century

MARGARET R. MEYER
Russell Sage College

"HOW CAN MOTHERS be educated and the nation made to recognize the supreme importance of the child?" [1]* This is a question from the notebook of Mrs. Alice McLellan Birney, under whose leadership the first Congress of Mothers was called. The date was February 17, 1897, but the problem was far from new, for there is arresting evidence that the germ of the idea was alive in the 1830's, and grew steadily through the next several decades. It appears repeatedly in the proceedings and reports of educational meetings and institutes, in contemporary journals and in the writings and speeches of such outstanding educators as Emma Willard and Henry Barnard. It is patent that there was widespread preoccupation with such perplexities as parental understanding of school problems and home-school relationships many years before the Parent Teacher movement became an organized activity in the late nineties.

Henry Barnard in 1839, when he called the first meeting in Connecticut to "further the cause of common schools," [2] held as part of his program a course of evening lectures to interest parents and the public. and later in 1847 planned discussion of "reciprocal duties of parent and teacher."[3] This idea of an evening meeting for the public persisted. It reappears in 1844 in New York when Monticello citizens at the Sullivan County Teachers' Institute held a public gathering to express their interest in the Institute and its workings. [4] The following year a Chatauqua County Institute asked itself "what is the best method of interesting parents?" The answer seems to have been one entire evening given over to a discussion of how to "awaken the spirit of inquiry in children." Could this be aided by teachers visiting parents or parents visiting schools? [5] In 1845 Madame Emma Willard, ever mindful of the cause of women, offered some resolutions at the Oneida County Institute which were debated and passed. Included were the forthright statements that women must be regarded as having high responsibility in the care of children; that this care should extend from the nursery to the village school. There was a concluding statement somewhat more subtly worded that if men "whom we recognize as by the laws of God and man, our directors, and to whose superior wisdom we naturally look for guidance," should call women into the field of active labor for common schools, they would "obey with alacrity, and do the tasks they judge suitable for us." [6]

On another occasion she remarked that women had for 300 years helped the clergy in Christian philanthropy, could stretch money better than men, and had a reserve power which, if channeled into the cause of schools, would insure success. [7]

As late as September 1854, instructions for programing institutes included holding lectures on the relation of schools to society, parents and teachers. [8] Three months after, the Montgomery County Institute emphatically resolved that the "cooperation of parents with teachers in the performance of their duties is essential to the prosperity of the schools." [9]

What did these earnest men and women expect from their early day advocacy of close cooperation between parent and teacher? In their own words, they hoped first to establish mutual sincerity, confidence and respect. Batchelder, in a plea before educators and laymen, said that parents should "strive to render the teacher's desk a post of honor and profit." [10] They saw, too, opportunities for cultivating correct social habits in the child's character; [11] and certain formally organized groups, such as the Connecticut Ladies of Kensington, anticipated improved physical conditions for the schools through the efforts of an "accommodation committee." Child, teacher, parent, and the educational system were to profit by the relationship.

A variety of problems was attacked in a multitude of ways. Foremost was the selection of the teacher. Although this was usually done by a committee, Abbott felt that parents must show some interest in the process, so that the committee would be “animated and encouraged,”[12] and the teacher would know that the community attached importance to his labors. Crowding teacher selection for priority was the question of good pay and its beneficial consequences. Abbott thought that no “money was so well expended as that which was made really to tell on the moral and intellectual improvement of children,”[13] and Batchelder remarked that a “teacher who at fifty dollars per month will secure the respect and obedience of his pupils, might utterly fail at the meagre price of half that sum.”[14] Even the Connecticut Ladies of Kensington concerned themselves, not only with school housekeeping, but with getting pay for extra teachers.[15]

---

* Footnotes will be found at end of article.

These same ladies had a second group called a “location committee” which grappled with the perplexing question of attendance, trying to see to it that every child of proper age was in school. Poor attendance was apparently intensified by the indifference of parents in the underprivileged group whose children, according to Batchelder, were never in school except on stormy or unemployed days, and who had no appreciation of the power of education to elevate their families from their forlorn conditions.[16]

Moreover, there was apprehension over the supervision of so-called leisure hours. Batchelder said parents must not expect the school to educate the whole child; that his time out of school was his parents' to control, and that study at home must be superintended. His pleas were eloquently concluded by the question, “If not the parent, who shall support the suppliant hand of the teacher raised in invocation. . . to aid him in the conflict with besieging hosts which distract the child?”[17]

But control of leisure time meant more to educators than mere supervision of home work. They believed that there was a deep-rooted connection between children's moral and social habits formed at home, and their intellectual progress at school, and parents in general were neglectful of this aspect of home preparation. Obedience, truth and reverence for authority were learned at home, and their teaching threw no heavy burden on fathers and mothers. Jason Whitman, in a speech before the American Institute of Instruction, claimed that teachers and parents were “too much estranged from each other,” tended to “look at opposite sides of the picture and thwart each other's efforts,”[18] while Abbott pointedly remarked that children “well trained at the fireside” made no trouble for their teachers.[19]

Construction and internal arrangement of schoolhouses were considered projects for parental activity. As mentioned above, the industrious Connecticut Ladies expected their “accommodation committee” to look after schoolhouses and grounds, and Henry Barnard's well-attended 1847 institute in Connecticut, offering public lectures on this topic, further included administration and management of schools among laymen's duties.[20]

A few speakers thought that the “literature of the school books” owned by the schools should be purchased through parental consultation, and had a special committee for this piece of work.[21]

With such a host of problems, one is forced to look for the measures used toward their solution. Public evening meetings at institutes, and societies similar to the one in Connecticut were fairly common. Much weight was placed on building up teacher-parent friendliness in communication about the child's needs and character, and to this end mutual visiting seems to have been widely urged. Batchelder recommended that the teacher make evening calls in the home, with children present, as the guest was “infused in the domestic circle.”[22]

About the reverse of such visiting, that which took fathers into the schools, much advice was given, with numerous warnings against possible abuses. Parental presence on examination days was quite frequent, but Batchelder argued that this custom gave only a partial view of school activities, and that attendance on other days afforded a better picture.[23] Committees of a trustee nature were used as means of checking-up and supervising, and Thomas Greene, who lectured enthusiastically on this topic, recommended that the average group be enlarged and made to take its responsibilities more seriously. The parent, he said, should “let the compting room be left. . . in charge of his partner or clerk. . . while he is for the present neither the merchant or professional man. Let him give full attention (to the school) and not look ever and again at his watch.”[24]

Apprehension lest the results of these visits be negative shows in numerous cautions for a constructive approach. Abbott warned against judging teachers on children's testimony, and on condemning them in their students' presence.[25] Interference with the curriculum was alarming to Greene,[26] and Batchelder said that if parents revealed to their friends anything amiss in the teacher's management of the school, the situation became worse as a result of gossip. He advised a private approach or consultation if, in a father's opinion, something was wrong,[27] and urged that all that was commendable be freely praised.[28]

Checking contemporary judgments of these early efforts of layman and professional to work together, we find Sweet reporting that the effect of the teacher institutes on the public was satisfactory. People attended evening lectures, were enlightened on many subjects, and no longer regarded the teacher with indifference.[29]

Teachers, on the other hand, through mutual visiting and the diverse work of committees, knew that the community was aware of them, and not indifferent to the importance of their work. Perhaps Batchelder, in his customary flowery phrases, best expressed the meaning of the whole movement when he said "where there is a good teacher, the parent has no cause to complain. . . . If the parent does his part, the Arctic regions of frigidity (between the two groups) will be thawed into mildness. Mutual respect and confidence will be established and will shed their benign influence on the spirit of the child."[30]

FOOTNOTES

1. Monroe, Walter S., ed. Encyclopedia of Educational Research (New York, 1950), p. 804.
2. American Journal of Education, XV (September 1865), p. 394.
3. Ibid.
4. Teachers' Advocate, I (October 29, 1854), pp. 120-121.
5. Sweet, Samuel N. Teachers' Institutes, or Temporary Normal Schools, Their Origin and Progress. . . (Utica, 1848), p. 86.
6. Ibid., p. 96.
7. District School Journal, VI (September, 1845), p. 117.
8. New York Teacher, II (September, 1854), p. 376.
9. Ibid., III (December, 1854), p. 145.
10. Batchelder, Jacob. "The Cooperation of Parents and Teachers," American Institute of Instruction, Journal of Proceedings, XIX (1848), p. 27.
11. Ibid., p. 43.
12. Abbott, Jacob. "The Duties of Parents in Regard to the Schools Where Their Children are Instructed," American Institute of Instruction, Lectures and Proceedings, V (1834), p. 83.
13. Ibid., p. 86.
14. Batchelder, op. cit., p. 29.
15. District School Journal, op. cit.
16. Batchelder, op. cit., p. 61.
17. Ibid., p. 38.
18. Whitman, Jason. "Home Preparation for School," American Institute of Instruction, Journal of Proceedings, XVII (1846), pp. 34-36.
19. Abbott, op. cit., p. 96.
20. American Journal of Education, op. cit.
21. District School Journal, op. cit.
22. Batchelder, op. cit., p. 43.
23. Ibid., p. 54.
24. Greene. Thomas A. "On the Duty of Visiting Schools," American Institute of Instruction, Journal of Proceedings, XI (1840), p. 53.
25. Abbott, op. cit., p. 96.
26. Green, op. cit., p. 62.
27. Batchelder, op. cit., p. 61.
28. Ibid., p. 54.
29. Sweet, op. cit., p. 28.
30. Batchelder, op. cit., p. 32.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 1, September 1962)

Effectiveness of Different Methods of Study

MABEL S. NOALL
Boston University

WHICH METHOD of study contributes most to high school and college students'learning and retention of facts— listening, listening with note-taking, reading, or reading with note-taking? This paper presents the findings from two studies, one on university students (3) and the other on high school students (2), with a survey test on a different university population (1) to check results of the first study.

Method

The first study used seventy-three students from the lowest achievement quartile of a university freshman class. The second study used a bright group of seventy-two tenth grade students. Since the two studies used the same research design and virtually the same materials, and since all students participated in all methods of study and used all materials, a number of interesting comparisons were possible without equation of groups. The survey testing used to check on reading and study skills used typical college freshmen at a university in a different part of the country to see if college students' performance in a different setting would agree with the findings of the first study. Although the three groups studied were all above average in intelligence, very good (2), average (1), and below average (3) students were included.

Material

Three (of four) exercises were used in each study and in the survey test. All of the exercises were comparable in average sentence length, average syllable count in a hundred words, number of major ideas presented, abstraction level of ideas, number of words in the article, and format. Three of the exercises were taken from Ruth Strang, Study Type Reading Exercises (4): Exercise 2, pp. 8-11; Exercise 9, pp. 45-48; Exercise 18, pp. 94-98. The fourth exercise, "How to Learn from a Lecture," was original, written to conform with the Strang exercises in style, length, content, and structure.

Procedure

Each week one article was presented to each of three classes with the method of presentation differing on a rotated basis. The time allowance was kept constant at four minutes. Immediately after the presentation, comprehension was tested. The following week a retest of the same material was given. Then a new article was presented by a different method and comprehension tested. After another week, again, there was a delayed recall of the material from the previous week; another article was presented and comprehension checked. This procedure was continued until all the materials by all three methods had been presented. Finally, a test of fifteen multiple-choice items, five on each of the three articles, was given to all the students. In the study with the college students, unaided recall with blank outlines was used in the weekly tests, but in the study with the high school students, aided recall tests of multiple-choice items were used.

Scoring

Scoring was done by idea count. Little or no inferential thinking was required. Organizational patterning was observed for confusion in relationships on the college group, but no penalties were applied for incorrect organizational structure.

Analysis of the Findings

By combining the scores of all students together for each specific method of presentation, means, standard deviations, and sigmas of the means were computed and critical ratios determined to check the significance of the differences. This was done on both the immediate recall and delayed recall tests, by sex and by total groups, and also on the final multiple-choice test. In this way the various kinds of methods were tested for relative effectiveness in delayed versus immediate recall, for relative effectiveness with boys versus girls, and for effect of recency on recall.

Results

(A) The college freshman study using listening versus listening with note-taking versus reading will be presented first. The group averaged 108 in I. Q. on the CAS Scale of the Wechsler-Bellevue Form I, 8.4 grade level on the Cooperative Reading Test C2T, 19 years of age, a predicted-grade-point average of 1.5 (D+), and had twice as many men as women in the group. Complete test results were available on 65 students, 45 men and 20 women. To facilitate intercomparisons, eight additional students on which only partial results were available were excluded from the study. Only relationships found to be significant at the one percent level of confidence are reported. Fisher's T tables were used to determine the level of significance after critical ratios were established. Results were as follows:

1. On immediate recall, reading was superior to listening with note-taking and was the only significant difference which existed among the three methods on immediate recall tests.
2. On delayed recall, reading was superior to listening, and it was also superior to listening with note-taking.
3. Comparing immediate and delayed recall on specific methods, immediate recall was significantly superior to delayed recall on both listening and listening with note-taking.
4. When reading was used there was practically no loss of learning on delayed recall tests given after one week lapse of time.
5. The women were superior to the men in immediate recall following both listening and listening with note-taking.
6. Recency of learning affected recall with the most recently presented information being best remembered and the scores dropping lower with longer time delays in the recall. The greatest loss occured on material learned more than three weeks before.
7. Students who recalled in organized fashion tended to score higher both on immediate and delayed recall tests.
8. Students scored significantly higher on multiple-choice tests which stimulated recall (80 percent) than on unaided outline recall (56 percent). In other words, the type of test given makes a difference.

(B) The second study was done with tenth graders using listening versus reading versus reading with note-taking. The group averaged 122 I. Q. on the Terman-McNamar Test, 15.5 years of age, and was about equally divided as to sex. Results were as follows:

1. On immediate recall all methods appeared to be about equally successful.
2. For the boys, reading and reading with note-taking gave the best learning and highest delayed recall.
3. For the girls, reading with note-taking gave the best results.
4. On reading with note-taking, the number of items answered correctly on delayed recall after one week was 5.7 percent greater than on the immediate recall test.
5. Reading with note-taking and reading were both superior to listening on delayed recall.

(C) Survey Test. Using the same reading and test materials for a survey of reading and study skills of 79 college freshmen at a different university, the mean scores were within .5 of the mean scores made by the experimental group used in the first reported study. This was true on immediate and delayed recall tests. Because all materials were presented and tested one week and all retests were given the following week, no detailed analysis of data was made.

Discussion

In interpreting results, caution must be observed in several areas. First, the materials used were very easy study-type materials of about fifth-grade difficulty, and the results obtained might not apply on difficult materials and under other conditions. Secondly, the close approximation of test means on the first group of university freshmen and the second group of university freshmen were obtained on tests of only fifteen items. Obviously, the close approximation of test means is partly a result of the brevity of the tests and probably would not have held if the tests had been longer. In the third place, the outline of a student who recalled a number of items was apt to appear better organized than the outline of the student who recalled fewer items. For this reason, the appearance of relationship between organization and better recall may be an illusion, and it does need further study.

REFERENCES

1. Hogan, Florence P., Johnson, Wendel, and Johnston, Thesba, Effectiveness of Recall Using Different Impression Techniques, unpublished research, Boston University, 1959.
2. Marr, Lucy U., The Comparative Value of Listening, Reading, and Reading with Note-Taking on Learning and Recall in Grade Ten, unpublished Ed. M. thesis, Boston University, 1958.
3. Noall, Mabel S., Listening Versus Listening with Note-Taking Versus Reading, unpublished research completed at the University of Utah in the spring of 1956.
4. Strang, Ruth, Study Type of Reading Exercises, Revised, New York: Bureau of Publications, Teachers College, Columbia, 1956.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 1, September 1962)

The Validity of Early Entrance into Kindergarten

ROSALEE G. WEISS

Menlo Park, New Jersey [1]

A CONTROVERSIAL issue in recent years has been the need for flexibility in entrance age within the school system and its relation to individual differences among children. The issue is stimulated by the increasing awareness concerning developmental differences among children, the rate at which they learn, and the degree of learning they can attain. Whereas current admission policies vary, admission by chronological age is typical. It is also known that many school districts are quite dissatisfied with their present admission policies, but are hesitant to change them due to the citizens' tendency to maintain the status quo.

Various acceleration methods are being used to admit pupils on the basis of individual differences. Among these, early entrance through testing programs has become quite popular throughout the country. Some schools will now permit children to enter kindergarten at an early age providing they are of superior intelligence, have an adequate degree of social and emotional maturity, and are considered to be in good physical health.

However, when a child is put in a classroom at an early age it might not be in the best interest of the child's achievement and adjustment. A child with a rapid rate of mental growth appears to be ready for school long before his chronological age, but is he really ready?

Research to compare the achievement and adjustment of bright and mature early-age children with normal-age kindergarten children under experimental conditions has been relatively neglected. Most of the literature concerning the problem of early-age entrance is based on educators' and clinicians' judgments, experiences and theories that may or may not be applicable for a given situation. None of these theories have been adequately tested.

On the basis of the above concern, the author[2] conducted an investigation concerning entrance age. The study involved the evaluation of the achievement and adjustment of selected emotionally and socially mature children with an I. Q. of at least one standard deviation above the norm in the Stanford Binet test who were admitted at an early age, by a testing program, into kindergarten classes specifically conducted for normal-age children.

The investigation was carried out experimentally by a comparison of the achievement and adjustment of 35 selected early-age children with 161 normal-age children with whom they had been matched. The matching factors included I. Q., personality adjustment and sex for children of above-average I. Q., and personality adjustment and sex for children of average and below-average I. Q.. The tests administered to match the subjects were the California Test of Mental Maturity and the California Test of Personality (individually administered).

At the end of the school year the achievement and adjustment of the control and experimental groups were compared by using the scores obtained from the California Test of Personality, a sociometric test and the Teacher Rating Scale. The instruments were administered individually.

The results of the mean performance scores at the end of the year indicated that the early-age children performed lower on rating scale grades, personality scores and social status than the normal-age children of comparable I. Q. and who had matching personality scores at the beginning of the school year. The early-age children with above-average I. Q. also had lower scores on these criterion variables than the children of average I. Q. who were 8 to 12 months older.

The following conclusions are based on the findings of this study.

1. When placed in a regular kindergarten class, early-age children of above-average I. Q. may be expected to achieve and adjust approxi-

---

[1] State of New Jersey Department of Institutions and Agencies, Diagnostic Center, Menlo Park, New Jersey

[2] Rosalee G. Weiss, The Validity of Early Entrance into Kindergarten. Unpublished Ph.D. Thesis, New York University, 1960.

mately at the level of the class average.

2. While the early-age children achieve and adjust better than the normal-age children of below-average I. Q., they may be expected to achieve below the level of children of comparable I. Q. who enter kindergarten as the oldest group.

These findings may be interpreted to mean that if the early-age children had entered kindergarten a year later in the normal pattern, they would have achieved at a higher level.

In interpreting these findings, the effect of average performance on the motivation and adjustment of superior early-age children should be considered.

Motivation is an important factor to consider in educational theory as it has a directing and reinforcing effect upon learning. When younger children are matched intellectually, emotionally and socially with older children, older children should have the higher level of maturation. Therefore, during the early years the younger children will have an age handicap in competing for achievement and social status in the classroom. Studies in readiness reveal that children learn with more zest and in less time if they are more mature. Research indicates that if a child is exposed to a situation before he is ready or capable of making sufficient progress for self-satisfaction, he may develop unfavorable attitudes toward this situation.

A child becomes interested in activities when he is successful and and when the activities are geared to his level of ability. When a child does not achieve at his level of ability, he tends to withdraw, with a resultant lowering in interests and motivation. Children tend to develop best with early experiences in realistic goal-setting and a sense of achievement.

As a rule, satisfaction is an index of success. Consequently, the bigger the reward or the satisfaction the more desired the incentive, and the reinforcement of incentives results in better performance. If superior early-age children continue to perform more poorly than superior normal-age children, their poorer grades tend to reduce motivation as they do not gain the success and approval due them.

Another important implication of the findings is in the area of group recognition.The regard accorded another individual by other members of his group exerts an important influence upon his behavior. Group recognition is a powerful incentive and one which the individual strives to attain in many ways, particularly through conformity and competition. In the findings of this study, the older normal-age children of average I. Q. and all children with above-average I. Q. were more popular at the end of the school year than the superior I. Q. early-age children. If this trend continues into later grades, early-age children will be at a disadvantage socially and in leadership opportunities.

If the pattern demonstrated in this investigation continues throughout the children's schooling, it must be decided whether superior I. Q. children should be exposed to experiences that may hinder rather than help their intellectual, emotional and social development.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 1, September 1962)

Selected Characteristics of School Board Members and Their Attitudes Toward Criticisms of the Public Schools

FRANK R. ALBERT, JR.

Inter American University of Puerto Rico

Introduction

THE RUSSIAN launching of the earth's first artificial satellite, Sputnik I, on October 4, 1957, marked the beginning of a new era of reawakened public interest in American public school education. Since its launching, the American people have been seriously concerned with the position of the United States in the race for the conquest of outer space. Criticisms of the public schools, though not new, have flourished in this period of rapid technological change and scientific discovery. Frequent re-examination of the programs of the public schools is both desirable and necessary if the needs of American youth and democracy are to be effectively served. While many of the criticisms of the public schools have stemmed from a sincere concern and interest in the welfare of American youth, motives underlying other criticisms, however, may be of a more questionable nature.

It was the purpose of the present study to determine the relationship between selected characteristics of public school board members of United States cities of 30,000 or more in population and their attitude toward selected criticisms of public school education. A secondary purpose of this study was to determine to what extent changes occurred in the composition of public school boards serving in cities of 30,000 or more in population as compared with the composition of school boards serving in cities of similar size in 1945.

Procedure

A two-part questionnaire was used in collecting data in this study. Part I consisted of items concerning the characteristics of the board members which were studied. These characteristics were age, sex, education, occupation, children, wards or grandchildren attending public schools, tenure, and income. In addition, replies were identified by the geographic region and city size in which the board members served. Typical of the criticisms used in Part II of the questionnaire were the following:

"5. There is a need for more rigid discipline in the public schools."
"15. The school door has been slammed in the faces of college graduates who want to teach because they have not taken superfluous courses in teacher education."
"20. Subjects and subject matter are being ignored in the schools and, instead, schools are teaching children."
"26. The schools are failing to teach children the fundamentals of the three R's."

Response categories of: "Strongly Agree," "Agree," "Undecided," "Disagree," and "Strongly Disagree" were assigned to each of the statements used in Part II of the questionnaire. The questionnaires were mailed to public school boards in 396 cities of 30,000 or more in population. The 727 replies used in the study represented an estimated twenty-seven per cent of an estimated 2,688 board members serving in cities of this size.

The data concerning the characteristics of board members obtained from Part I of the questionnaire were compared with the findings of a related study of school board composition which was conducted in 1945.* The analysis of variance technique and the t test were used in examining the relationships between the characteristics of board members and their attitudes toward criticisms contained in Part II of the questionnaire. A significance level of one per cent was used in the present study.

---

*Research Division of the National Education Division, "Status and Practices of Boards of Education," Research Bulletin, XXIV, no. 2 (Washington: National Education Association, 1946).

Results

Board members differed significantly in their attitudes toward criticisms of the public schools when the scores attained on Part II of the questionnaire were classified by geographic region, age and sex. Board members serving in the Middle Atlantic States indicated more agreement with negative criticisms of the public schools than board members serving in the Middle States and Far West region. Board members in the Middle States indicated less agreement with criticisms of the schools than those serving in the Southeast region. Board members over sixty years of age showed significantly greater agreement with the criticism than did board members from forty to fifty years of age. Women serving on school boards indicated less agreement with the statements of criticism than did male board members. The median age of school board members in cities of 30,000 or more in population was 48.6 years in 1959 as compared with 50.8 years in 1945. Representation of women on the board membership increased from fourteen per cent in 1945 to eighteen per cent in 1959. Professionals, proprietors and executives were the occupational groups most frequently represented on the school boards in both 1945 and 1959. Seventy-two per cent of the board members in 1959 were college graduates, twenty-four per cent were high school graduates, and the formal education of four per cent ended before high school graduation. Seventy per cent of the board members in 1959 had one or more children attending public schools. Fifty-seven per cent of the board members had served from two to nine years on school boards. The median period of service was six years. The median annual income of board members was $11, 968.

Conclusions

The attitudes of board members toward criticisms of the public schools do not seem to be significantly related to board members' occupations, educational backgrounds, length of service on the school board, income, the city size in which the board members serve, or to whether or not the board members have children, wards, or grandchildren attending public schools. Board members most likely to agree with current criticism of the public schools would be board members over sixty years of age and board members of either sex serving in the Middle Atlantic and Southeast States. The typical public school board member in cities over 30,000 in population is a man in his late forties; is married and has children, wards, or grandchildren attending public schools; is a proprietor, executive or professional who has an annual income of some $11, 968; is a college graduate; and has served six years on the school board. Approximately two out of every ten board members are women. The typical public school board in 1959 as compared with the typical board in 1945, had a larger representation of women and was comprised of younger board members, more of whom were married with children attending the public schools. The board members in 1959 followed approximately the same vocations as did the 1945 board members, had a higher economic and educational status, and served approximately the same period of service on the school board.

REFERENCES

1. Benjamin, Harold, and others. " The Attacks on Modern Education, " in *Freedom and Public Education*, Ernest O. Melby, editor (New York: Frederick A. Praeger, Inc., 1953), pp. 135-216.
2. Champlin, Nathaniel L. " The Attacks Upon Education: Their Significance for Philosophy of Education, " *Educational Theory*, VIII (July, 1958), pp. 157-161.
3. Counts, George S. "The Composition of Boards of Education." *Supplementary Educational Monographs*, XXXIII (July, 1927).
4. Gross, Neal. *Who Runs Our Schools?* (New York: John Wiley and Sons, 1958).
5. Research Division of the National Education Association. "Status and Practices of Boards of Education, " *Research Bulletin*, XXIV (April 1946), pp. 47-83.
6. Research Division of the National Education Association. " Ten Criticisms of Public Education, " *Research Bulletin*, XXXV (December, 1957), pp. 141-148.

Carbondale, Illinois. Critique Editor.
WILLIAM A. BROWNELL, Dean, School of Education, University of California, Berkeley, California. Education Psychology, Learning, and Teaching.
ROBERT A. DAVIS, Professor of Educational Research, George Peabody College for Teachers, Nashville, Tennessee, Evaluative Techniques.
GUY D. McGRATH, Dean, College of Education, Arizona State University, Tempe, Arizona. Secondary Education.
HAROLD E. MITZEL, Director, Office of Research and Evaluation, Division of Teacher Education,. College of the City of New York, New York City. Teacher Personnel.
WILLARD C. OLSON, Dean of College Education, University of Michigan, Ann Arbor, Michigan. Child Development.
JOHN SCHMID, Professor, College of Education, University of Arkansas, Fayetteville, Arkansas. Field News Editor.
WILSON B. THIEDE, Professor of Education, University of Wisconsin, Madison, Wisconsin. Book Review Editor.
ARTHUR E. TRAXLER, Executive Director, Educational Records Bureau, New York 32, New York. Educational Measurements.

Associate Editors

HAROLD M. ANDERSON, Associate Professor of Education, University of Colorado, Boulder, Colorado.
VERNON E. ANDERSON, Dean, College of Education, University of Maryland, College Park, Maryland.
RUDYARD K. BENT, Professor of Education, University of Arkansas, Fayetteville, Arkansas.
EMMETT ALBERT BETTS, Research Professor, University of Miami, Coral Gables, Florida
LAWRENCE P. BLUM, Assistant Director, Student Personnel Services, University of Wisconsin–Milwaukee, Milwaukee 3, Wisconsin.
J. MURRAY LEE, Chairman, Department of Elementary Education, Southern Illinois University, Carbondale, Ill.
D. WELTY LEFEVER, Professor of Education, University of Southern California, Los Angeles, California.
L. JOSEPH LINS, Professor and Coordinator, Institutional Studies, University of Wisconsin, Madison, Wisconsin.
S. E. TORSTEN LUND, Associate Professor of Education, University of California, Berkeley, California.
WILLIAM A. McCALL, Lecturer in Education, University of Miami, Coral Gables, Florida.
DELYTE W. MORRIS, President, Southern Illinois University, Carbondale, Illinois.
VALWORTH PLUMB, Chairman, Division of Education and Psychology, University of Minnesota (Duluth Branch), Duluth, Minnesota.
DAVID G. RYANS, Rand Corporation, Santa Monica, California.
LESTER B. SANDS, Professor of Education, University of California, Santa Barbara, Goleta, California.
MANDEL SHERMAN, President, and Consultant, Emotional Health Association, Los Angeles, California.
MILDRED B. STANTON, Consultant, Special Education, State Department of Education, Hartford 15, Conn.
W. W. THEISEN, Professor of Education, Marquette University, Milwaukee, Wisconsin.
M. R. TRABUE, Professor of Higher Education, University of Kentucky, Lexington, Kentucky.
J. WAYNE WRIGHTSTONE, Director, Bureau of Educational Research, Board of Education of City of New York, Brooklyn, New York.

Editors Emeriti

E. J. ASHBAUGH, Dean Emeritus, School of Education, Miami University, Oxford MR50, Ohio.
HARRY J. BAKER, 19050 Wiltshire, Lathrup Village, Michigan.
FOWLER D. BROOKS, Director Emeritus, Department of Psychology, DePauw University, Greencastle, Ind.
LEO J. BRUECKNER, Professor Emeritus, University of Minnesota, Minneapolis, Minnesota.
B. R. BUCKINGHAM, 33 Chatham Circle, Wellesley Hills, Massachusetts.
EDGAR A. DOLL, Director Emeritus, Devereux Schools, Devon, Pennsylvania.
HARL R. DOUGLASS, Director Emeritus, School of Education, University of Colorado, Boulder, Colorado.
E. DUNCAN GRIZZELL, Dean Emeritus, University of Pennsylvania, 640 Maxwelton Court, Lexington, Ky.
ERNEST HORN, Professor Emeritus of Education, State University of Iowa, W307 East Hall, Iowa City, Iowa.
EARL HUDELSON, Professor Emeritus of Education, West Virginia University, Morgantown, West Virginia.
A. M. JORDAN, Professor Emeritus, University of North Carolina, Chapel Hill, North Carolina.
ALFRED S. LEWERENZ, Director Emeritus of Evaluation and Research, 4107 Holly Knoll Drive, Los Angeles, California.
A. R. MEAD, Director Emeritus of Educational Research, University of Florida, 1719 N. W. 6th Ave., Gainesville, Florida.
M. J. NELSON, Dean Emeritus, Iowa State Teachers College, 6033 West Orange Drive, Glendale, Arizona.
JESSE B. SEARS, Professor Emeritus, Stanford University, Stanford, California.
CARLETON W. WASHBURNE, Chairman Emeritus, Department of Education, Brooklyn College, Brooklyn, New York.

Staff

JEAN WILLIAMS, Supervisor of Editorial Services. **CLARENCE A. SCHOENFELD**, Editorial Associate. **ARNOLD CAUCUTT**, Business Manager.

The Journal of Educational Research

Dedicated to the Scientific Study of Education

VOLUME 56 **OCTOBER 1962** **NUMBER 2**

It Says in JER 58

Among Those Present 60

The Relationship of First-Grade Readiness to Third- and Fourth-Grade Achievement—Albert J. Kingston 61

Likes and Dislikes of Indonesian Pupils—R. Murray Thomas and Winarno Surachmad 68

A Study of Morale in Education Utilizing Incomplete Sentence—John H. Suehr 75

Teacher Time Utilization with Teacher Aides—Curt Stafford 82

Factors Involved in the Adjustment of Children Entering First Grade—Sarah Lou Hammond and Dora S. Skipper 89

Value Variations in Fourth- and Seventh-Grade Students—Phyllis Swartz, Edward G. Nolan, and Kenneth Tillman 96

The Development of a Parent Attitude Toward Education Scale—Gene R. Medinnus 100

Behavior Problems of Children as Viewed by Children in the Fifth through Eighth Grades—Thomas Mangan and David Shafer 104

The Laboratory School: An Historical Perspective—Pose M. Lamb 107

Field News—John Schmid 110

The Journal of Educational Research is published at Madison, Wisconsin, monthly 10 times a year (the May–June and July–August issues being combined) by Dembar Publications, Box 1605, Madison 3, Wisconsin. The subscription price is $7.50 a year plus $1 a year postage for all subscriptions outside the continental United States. Single copies, $1. Address all communications to Dembar. Second class postage paid at Madison, Wisconsin.

COPYRIGHT 1962 BY DEMBAR PUBLICATIONS, MADISON 3, WISCONSIN

It Says in the JER . . .

IN A STUDY OF VALUE VARIATIONS BETWEEN fourth and seventh graders, Swartz, Tillman, and Nolan found positive correlations for boys and girls in the fourth grade, boys and girls in the seventh grade, and between the total fourth and seventh-grade groups; but agreement among the fourth graders was more marked than was the case with the seventh graders. So the researchers postulate that value discrepancies based on sexual identification become more apparent with time.

KINGSTON CONFIRMS OTHER REPORTS that there is a positive relationship between first-grade readiness and later third- and fourth-grade achievement, and he suggests that certain of the sub-tests of the Metropolitan Readiness Tests are more valuable than others in predicting future scholastic achievement.

THE LABORATORY SCHOOL is no modern invention, Lamb reminds us. As early as the 17th century there were European institutions where teachers might "through practice, learn that for which they will, in the future, be employed." The first practice schools in the United States were privately owned and operated. By 1875, some 47 of the 67 state normal schools had laboratory schools attached. The leadership in educational research provided by Dewey and Judd at Chicago and Meriam at Missouri around the turn of the century is cited as being of great significance by most writers concerned with the development of the lab school.

WHEN GRADE-SCHOOL CHILDREN ARE ASKED to rate the seriousness of various behavior problems, their answers correlate highly with those of traditional teachers, with high-school seniors, and with college seniors; but there is less correlation with the views of modern teachers and a negative relationship to the opinions of mental hygienists. So report Mangan and Shafer in concluding that children tend to acquire the attitudes of their teachers toward behavioral patterns and retain these attitudes through high school and college.

WHAT FACTORS ARE INVOLVED in the adjustment of children entering first grade? Hammond and Skipper report a significant relationship between high-adjustment status and 1) attendance at kindergarten, 2) scores on reading readiness tests, 3) the socio-economic status of the family, and 4) chronological age. They found no significant correlations with size of family or ordinal position in the family.

WHAT HAPPENS IN THE CLASSROOM when teacher-aide programs are initiated as a partial solution to the problem of increased enrollment in the face of teacher shortage and inadequate facilities? Stafford reports on a detailed quantitative analysis of shifts in elementary teacher time utilization before-aide and after-aide. He discovers that 1) teacher activities of a clerical nature decreased significantly, 2) teacher activities requiring professional preparation failed to show a significant increase, and 3) the teacher alone was unable to maintain individual attention in an amount comparable to the pre-aide situation, but the aide herself seemed to share in this activity. He speculates that a qualitative change may take place in teacher performance in the presence of an aide, and suggests that the aide becomes an integral part of the classroom picture, not a mere appendage.

AN INSTRUMENT OF THE INCOMPLETE-SENTENCE TYPE has been developed by Suehr to measure teacher morale in terms of both environmental conditions and personality structure. After testing the device in Colorado, he believes it is the best method yet devised because it uncovers a broad area of trouble spots. Among Suehr's conclusions are these: 1) the childhood background of teachers is highly important to their morale as adult teachers, and 2) internal communication is probably the most vital area in the whole morale process. Hence Suehr suggests 1) there is too much haphazard encouragement of people to become teachers, and 2) a procedure for dealing with grievances should be operating in every school system.

ON THE THEORY THAT STUDENTS LEARN MORE EFFICIENTLY if they like their classwork, Thomas and Surachmad set out to study the likes and dislikes of pupils; in this case, some 1450 third through twelfth graders in West Java. In general, pupils expressed a greater total number of likes than dislikes, although there was a slightly greater variety of dislikes than likes expressed. The likes focused primarily on school subjects, the dislikes on distrubances in personal-social relations. The researchers speculate that since in many ways the pupils' answers reflect the same kinds of likes and dislikes adults might hope children would have, it may be that pupil expressions indicate what they think their teachers prefer rather than what the students themselves really believe. The authors also draw some conclusions as to why certain subjects show up strongly in both the “like” and “disliked” columns.

THAT HIS FAMILY'S ATTITUDES TOWARD EDUCATION can affect the child's performance in school has long been recognized. Now Medinnus reports the development of a reliable parent-attitude-towards-education scale that should

Continued on page 67. . . . .

Among Those Present

ALBERT J. KINGSTON, Jr., is an associate professor in the department of educational psychology, College of Education, University of Georgia. An acknowledged authority in the fields of reading, testing, and guidance, he has published widely. Dr. Kingston took his undergraduate work at the New York State Teachers College in Buffalo, his advanced degrees at Cornell. Previous professional experiences include director of remedial reading and guidance at Texas A and M College; remedial education advisor, U. S. Operations Mission to Ethiopia; and remedial specialist for the Atlanta public schools.

JOHN H. SUEHR, a Colorado EdD, is principal of the high school at Alamosa, Colorado. An Illinois native, he attended Western Illinois and Bradley and taught in Illinois before teaching in Colorado.

GENE R. MEDINNUS studied at the University of Minnesota and the University of Iowa, receiving the doctoral degree in child development from the Institute of Child Welfare, University of Minnesota. From 1956 to 1959 he conducted a three-year field study of school readiness and first-grade adjustment as a research associate with the Elizabeth McCormick Memorial Fund of Chicago. Since 1959 Dr. Medinnus has been on the psychology department staff of San Jose State College, California.

POSE M. LAMB is an assistant professor at Ball State Teachers College and a teacher in the Burris Laboratory School at Muncie, Indiana. Dr. Lamb holds three degrees from Ohio State and has taught in Michigan and Ohio.

SARAH LOU HAMMOND holds three degrees from Florida State where she is now professor of education. An international authority on elementary education, she has published widely and directed studies in Georgia and Texas as well as in her home state.

DORA SIKES SKIPPER is associate professor of education and director of off-campus instruction at Florida State. She has studied in Florida and at George Peabody, Chicago, and Columbia. A nationally known consultant, she was recently cited for her service to education in Florida.

PHILLIS SWARTZ was previously a graduate student at the University of New Mexico and is now employed by the Sandia Corporation at Albuquerque. EDWARD G. NOLAN is assistant professor of psychology at New Mexico, and KENNETH TILLMAN is a PhD candidate in physical education at UNM.

R. MURRAY THOMAS has AB and MA degrees from Colorado State College and a PhD in educational psychology from Stanford University. From 1958 through mid-1961 he was director of the State University of New York teacher-education unit at Bandung, Indonesia. He is now a professor of education at the University of California, Santa Barbara.

WINARNO SURACHMAD was born in South Celebes and schooled under the Dutch-colonial, Japanese-occupation, and Indonesian-revolutionary regimes. He earned his AB at Gadja Mada University in Central Java. After receiving an MA from the State University of New York, College at New Paltz, in 1958 he taught for three years at Padjadjaran University, Bandung. He recently began doctoral work in educational psychology at Stanford University.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 2, October 1962)

The Relationship of First-Grade Readiness to Third- and Fourth-Grade Achievement

ALBERT J. KINGSTON, Jr.
University of Georgia

THE CONCEPT OF READINESS has gained wide acceptance among elementary school teachers and administrators. While the term is not definitive, and often includes a wide range of characteristics, attributes, skills, and abilities, generally the concept includes any of the characteristics of a pupil which are deemed to contribute to his ability to profit from instruction. Although the concept is utilized at all levels of the school program, first-grade teachers traditionally are most concerned with methods of appraising the child's readiness to profit from school experiences. A number of standardized tests have been developed to assist in appraising the readiness of beginning first graders. Such measures have been shown to have some value in differentiating entering first-grade pupils.

The Problem

The present study was undertaken to ascertain the relationship of beginning first-grade readiness test scores to third- and fourth-grade scholastic achievement. School achievement results from a combination of many factors and conditions. Mental ability, physical and mental health, pupil attitude and motivation, developmental rate, the caliber of instruction and type of instructional materials employed are only a few of the factors which contribute to academic achievement.

Despite the many variables which are involved in the learning process, it is possible that children who show the greatest readiness for learning at the first-grade level would tend to make greater scholastic achievements at the third- and fourth-grade levels. Research in other areas of growth and development, on the other hand, show that many slower-developing children "catch up" during the elementary school years.

The relationship of first-grade readiness and third- and fourth-grade achievement as measured by well-known standardized tests was investigated by means of multiple regression equations.

The Instruments

The instruments employed in this investigation are widely known tests. The Metropolitan Readiness Tests (5) are probably the most commonly used tests of first-grade readiness. They consist of six subtests contained in a single booklet. The tests are: word meaning, sentences, information, matching, numbers, and copying. The tests themselves consist of pictures or figures which the child either marks or copies according to instructions administered orally. A number of studies have shown that test scores attained at the beginning of the first grade correlate significantly with first- and second-grade achievement(3) (4). Form S was employed in the present study.

The Stanford Achievement Tests (6) are designed to appraise scholastic achievement in various subjects throughout the elementary grades. The primary battery is designed for use at the end of grade one, in grade two, and for the first half of grade three. It consists of a single booklet containing five tests: paragraph meaning, word meaning, spelling, arithmetic reasoning, and arithmetic computation. The elementary battery used in grade four of this study includes a language test in addition to the five contained in the primary battery. The tests have been extensively used and have undergone several revisions.

Procedure

Entering first-grade pupils in the Atlanta public schools are routinely administered the Metropolitan Readiness Tests during their first month of school. Third- and fourth-grade pupils are tested with the Stanford Achievement Tests during the month of September. The primary battery is administered at the third-grade level and the

TABLE I

RELATIONSHIP OF FIRST-GRADE METROPOLITAN READING READINESS TEST SCORES TO THIRD-GRADE STANFORD ACHIEVEMENT SCORES

(boys)

N = 134

Stanford Achievement Test (Primary)

Met. Scores	Para. Mean.		Word Mean.		Aver. Read.		Spelling		Arith. Reason.		Arith. Compu.		Aver. Arith.		Battery Median	
	beta	t^2	beta	t^2	beta	t^2	beta	t^2	beta	t^2	beta	t^2	beta	t^2	beta	t^2
Word Mean.	-.0358	.948	-.0122	.133	-.0242	.509	-.0426	1.633	.0138	.242	-.0244	2.447	-.0045	.060	-.0256	1.054
Sent. Mean.	-.0217	.430	-.0313	1.062	-.0293	.914	-.0256	.722	.0361	2.017	-.0179	1.662	.0085	.260	-.0108	.231
Inform.	.0533	1.412	.0019	.002	.0308	.554	.0584	2.060	.0202	.344	.0277	2.124	.0251	1.243	.0420	1.191
Matching	.0583	7.102**	.0437	4.740*	.0500	6.105*	.0345	3.024	.0315	3.538	.0181	3.803	.0248	5.077*	.0366	6.095*
Number	.0396	6.391*	.0379	7.000**	.0421	8.496**	.0381	7.219**	.0333	7.700**	.0068	1.071	.0203	6.681*	.0305	8.296**
Copying	.0279	1.459	.0395	3.476	.0292	1.860	.0142	.459	.0046	.0680	.0098	1.009	.0069	.352	.0189	1.452
Constant (k)	1.379		1.912		1.651		2.165		1.168		2.676		1.915		1.809	
R^2 =	.257		.233		.258		.182		.297		.103		.274		.279	

* Significant at 5% l. c.

** Significant at 1% l. c.

TABLE II

RELATIONSHIP OF FIRST-GRADE METROPOLITAN READING READINESS TEST SCORES TO THIRD-GRADE STANFORD ACHIEVEMENT SCORES

(girls)
N = 138

Stanford Achievement Test (Primary)

Met. Scores	Para. Mean.		Word Mean.		Aver. Read.		Spelling		Arith. Reason.		Arith. Compu.		Aver. Arith.		Battery Median	
	beta	t^2	beta	t^2	beta	t^2	beta	t^2	beta	t^2	beta	t^2	beta	t^2	beta	t^2
Word Mean.	.03838	.770	.0259	.420	.036	.840	.0099	.080	.0080	.075	.0024	.023	.0062	.0974	.0294	1.192
Sent. Mean.	-.009	.050	-.0093	.060	-.0069	.030	-.0476	2.04	-.0061	.046	.0068	.186	.0025	.016	.0002	.0000
Inform.	.03136	.420	.0459	1.12	.0385	.790	.0231	.380	.0273	.733	.0018	.016	.0129	.346	.0208	.500
Matching	.05732	5.340*	.0551	5.98*	.0527	5.52*	.0324	2.770	.0202	1.493	.0078	.730	.0131	1.328	.0359	5.546*
Number	.05509	6.630*	.0447	5.29*	.4987	6.640*	.0353	4.430*	.0378	7.044**	.0057	.522	.0231	5.580*	.0274	4.370*
Copying	.0269	.780	.0214	.600	.0221	.640	.0273	1.31	.0106	.273	.0252	4.97*	.0172	1.53	.02330	1.545
Constant (k)	.5087		.6975		.5905		2.1122		1.7143		2.3272		2.0217		1.2641	
R^2 =	.326		.317		.339		.190		.225		.141		.235		.332	

* Significant at 5% l. c.
** Significant at 1% l. c.

elementary battery at the fourth-grade.

A sample of 272 fourth- and fifth-grade pupils was selected from five different white elementary schools. Scores on the first-grade Metropolitan Readiness Tests, the third-grade Stanford Achievement Test and the fourth-grade Stanford Achievement Test were obtained from test copies filed in the pupils' cumulative records. The children and schools employed in the study were felt to be representative of the school system. Children whose records indicated poor health history, irregular attendance, or physical handicaps were not employed in the study.

Multiple regression equations were calculated to ascertain the relationships of total scores and subtest scores on the Metropolitan Readiness Tests and the total and subtest scores on the third- and fourth-grade Stanford Achievement Tests. Because of possible sex differences, the scores of boys were treated separately from girls. Only 269 pupils were employed in calculations of the fourth-grade scores because of incomplete test records.

Results

First-grade readiness test scores were found to correlate significantly with the battery median and subtest scores of the Stanford Achievement Tests at both levels. Significant relationships were also obtained between certain part scores of the Metropolitan and subtest scores of both Stanford Tests. These relationships are shown in the attached tables. Tables I and II show the beta weights, t^2 and r^2, obtained between each of the Metropolitan and the third-grade Stanford (Primary). For both sexes, the matching and number subtests of the Metropolitan correlated significantly with third-grade achievement. Matching correlates significantly with paragraph reading, word meaning, average reading, and battery median scores on the Stanford. The number test correlates significantly with the paragraph meaning, word meaning, average reading, spelling, arithmetic reasoning, average arithmetic, and battery median scores for both sexes. The only other significant correlation was found between copying and arithmetic computation.

Fourth-grade achievement as measured by the Stanford Achievement Tests - Elementary Battery also correlated significantly with first-grade Metropolitan scores. The most significant relationships between part scores was found with the number test of the Metropolitan. For girls, the number test correlated with all the Stanford Tests. For boys, significant correlations were found between number scores on the Metropolitan and all of the Stanford scores except arithmetic computation. In addition, copying and information scores were significantly related to certain Stanford part scores. These relationships can be seen in Tables III and IV.

Discussion

The present study tends to substantiate other research concerning the relationship of first-grade readiness to subsequent scholastic achievement. Bremer (1) found a correlation of +.40 between first-grade Metropolitan scores and beginning second-grade achievement. Karlin (4) correlated first-grade Metropolitan scores with a third-grade reading test and obtained a coefficient of +.36. With chronological age and intelligence factors removed, the coefficient dropped to +.25. Dean (3) found that Metropolitan Readiness Tests scores correlated +.59 with Metropolitan Achievement scores obtained six months later. In the present study, coefficients between Metropolitan Readiness Tests scores and third- and fourth-grade achievement ranged between +.3 and +.6.

The present investigation yields new information concerning the subtests comprising the Metropolitan Readiness Tests. At both third- and fourth-grade levels, the number and matching subtests correlated significantly. The matching test appears to be largely a test of visual perception. Children are requested to identify pictures of animals, objects, figures, or words which are similar. A number of previous investigations have shown that perceptual development and ability are positively related to reading achievement. Stroud (7) found significant relationships (around +.50), between rate of reading and rate of visual perception as measured by tests of word selection, letter perception, and number selection. Coleman (2) found that perceptual retardation was a significant factor in reading disability.

The number subtest of the Metropolitan appears to be heavily weighted with items which appraise intellectual qualities. These tests measure a variety of achievements including counting, number vocabulary, the recognition of written numbers, the interpretation of numerical symbols, etc. Many of the items undoubtedly reflect the child's rate of mental development. Mental ability, of course, is significantly related to scholastic achievement at all age and grade levels.

The results of this study would seem to imply that certain of the subtests of the Metropolitan Readiness Tests are more valuable than others in predicting future scholastic achievement. A number of implications for appraising readiness may be suggested. What validity do the subtests of the Metropolitan have in the prediction of scholastic growth? Will mental ability and visual perception tests achieve the same purpose in appraising readiness that the more common readiness tests do?

TABLE III

RELATIONSHIP OF FIRST-GRADE METROPOLITAN READING READINESS TEST SCORES TO FOURTH-GRADE STANFORD ACHIEVEMENT SCORES

(boys)
N = 133

Stanford Achievement Test (Intermediate)

Met. Scores	Para. Mean.		Word Mean.		Aver. Read.		Spelling		Language		Arith. Reason.		Arith. Compu.		Aver. Arith.		Battery Median	
	beta	t^2	beta	t^2	beta	t^2	beta	t^2	beta	t^2	beta	t^2	beta	t^2	beta	t^2	beta	t^2
Word Mean.	.0500	.799	-.0175	.119	.0150	.085	.0151	.128	.0505	.466	.0088	.007	.0304	1.550	.0213	.646	.0099	.077
Sent. Mean.	-.0045	.008	.0167	.133	.0042	.009	-.0403	1.120	-.0683	1.045	-.0109	.122	-.0153	.483	-.0143	.357	-.0251	.604
Inform.	.0538	.620	.1035	2.769	.0731	1.355	.0666	1.673	.0891	.974	.0471	1.237	.0007	.001	.0215	.444	.0689	2.485
Matching	.0330	.980	.0257	.719	.0349	1.299	-.0050	.041	.0107	.059	.0193	.088	.0376	6.696*	.0281	3.177	.0127	.356
Number	.0701	8.665**	.0473	4.752*	.0572	6.832*	.0461	6.581*	.1116	12.546**	.0510	11.944**	.0131	1.608	.0342	9.172**	.0527	11.985**
Copying	.0639	3.278	.0473	2.174	.0563	3.018	.0363	1.869	.0983	4.445*	.0254	1.355	.0293	3.635	.0260	2.426	.0454	4.061*
Constant (k)	.5654		1.308		.9953		2.305		-.0897		2.042		2.611		2.322		1.820	
R^2 =	.281		.214		.263		.145		.266		.260		.250		.296		.288	

* Significant at 5% l. c.
** Significant at 1% l. c.

TABLE IV

RELATIONSHIP OF FIRST-GRADE METROPOLITAN READING READINESS TEST SCORES TO FOURTH-GRADE STANFORD ACHIEVEMENT SCORES

(girls)
N = 136

Stanford Achievement Test (Intermediate)

Met. Scores	Para. Mean.		Word Mean.		Aver. Read.		Spelling		Language		Arith. Reason.		Arith. Compu.		Aver. Arith.		Battery Median	
	beta	t^2	beta	t^2	beta	t^2	beta	t^2	beta	t^2	beta	t^2	beta	t^2	beta	t^2	beta	t^2
Word Mean.	.0489	.784	.0254	.269	.0353	.508	.0100	.058	.0844	.298	-.0010	.0002	-.0045	.035	.0074	.075	.0202	.288
Sent. Mean.	.0106	.038	-.0052	.012	.0075	.025	-.0649	2.590	-.0206	.069	.0054	.022	.0089	.144	.0083	.099	.0058	.026
Inform.	.1181	3.683	.1182	4.854*	.1247	5.303*	.0352	.599	-.0375	.177	.0293	.0498	-.0042	.0249	.0062	.044	.0397	932
Matching	.0555	3.034	.0612	4.859*	.0557	3.956*	.0407	2.990	.1033	5.022*	.0302	1.985	.0161	1.370	.0175	1.292	.0475	4.975*
Number	.0723	6.921**	.0485	4.103*	.0528	4.774*	.0519	6.550*	.0977	6.045*	.0495	7.142**	.0262	4.899*	.0413	9.637**	.0523	8.092**
Copying	.0516	1.737	.0342	1.004	.0485	1.983	.0275	.909	.0731	1.666	.0360	1.867	.0316	3.488	.0316	2.772	.0412	2.475
Constant (k)	-.2982		.6193		.1843		2.729		1.113		2.281		3.244		2.710		1.587	
R^2 =	.390		.349		.388		.218		.271		.273		.218		.303		.383	

* Significant at 5% l. c.
** Significant at 1% l. c.

Summary and Conclusions

Multiple regression equations were employed to determine the relationships between first-grade Metropolitan Readiness Tests scores and third-grade Stanford Achievement Test (Primary) scores, and between first-grade Metropolitan and fourth-grade Stanford Achievement Test (elementary) scores for a group of elementary children from five public schools. The findings were:

1. First-grade Metropolitan Readiness scores correlated significantly with scholastic achievement in all areas measured by the Stanford Achievement Tests at both the third- and fourth-grade levels.
2. The magnitude of the coefficient obtained (between +.3 and +.6), indicate that the prediction of third- and fourth-grade achievements for individual pupils based upon their first-grade readiness scores is not feasible.
3. The matching and number tests of the Metropolitan Readiness Tests correlate with overall achievement at the third- and fourth-grade level to a greater degree than do the other parts of the Metropolitan. The matching test appears to be largely a measure of perception, while the number test seems to be heavily weighted with items which place a premium on mental ability.

REFERENCES

1. Bremer, Neville. "Do Readiness Tests Predict Success in Reading?" Elementary School Journal, LIX (1959), pp. 222-224.
2. Coleman, J.C. "Perceptual Retardation in Reading Disability Cases," Journal of Educational Psychology, XLIV (1953), pp. 497-503.
3. Dean, Charles. "Predicting First Grade Reading Achievement," Elementary School Journal, XXXIX (1939), pp. 609-616.
4. Karlin, R. "The Prediction of Reading Success and Reading Readiness Tests," Elementary English, XXXIV (1957), pp. 320-322.
5. Metropolitan Reading Tests. World Book Company (Yonkers-on-Hudson, 1950).
6. Stanford Achievement Tests. World Book Company (Yonkers-on-Hudson, 1953).
7. Stroud, J. B. "Rate of Visual Perception as a Factor in Rate of Reading," Journal of Educational Psychology, XXXVI (1945), pp. 487-498.

IT SAYS (continued) .....

facilitate research in the whole area. The scale consists of 40 statements, with each of which the parent is asked to agree or disagree. When he used the scale, Medinnus got some surprising results: 1) no statistically significant differences showed up among parents on the basis of education or socio-economic level, 2) the correlation between mothers' and fathers' scores was low, and 3) teacher assessments of parental attitudes correlated with actual father attitudes but not with mother attitudes. Medinnus suggests that use of his scale can be profitable in school administration after it has been further validated and refined.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 2, October 1962)

Likes and Dislikes of Indonesian Pupils*

R. MURRAY THOMAS
State University of New York
WINARNO SURACHMAD
Padjadjaran University, Indonesia

MOST EDUCATORS today seem convinced that students usually learn more efficiently if they like their class work than if they dislike it. That is, the pupil who enjoys the subject matter and the classroom climate will, in general, work more willingly to succeed in the class.

Hence it appears desirable for teachers to discover what their pupils like and dislike about school and school subjects so that measures may be taken to support the appropriate likes and to correct undesirable dislikes in ways that will carry pupils more surely toward the school's goals.

For each teacher this matter of investigating likes and dislikes is a personal responsibility to be carried out specifically for his own class. However, as he tries to evaluate the nature of his pupils' responses, it may help him obtain some perspective if he also knows opinions that have been expressed by pupils in other schools and at other grade levels. Thus he can answer: Are there some general trends of likes and dislikes that are more prominent at one age level than at another? Are the opinions expressed in my class unique or are they rather typical of other pupils at this grade level?

This study has had two main purposes:

1. To provide Indonesian teachers with a general picture of likes and dislikes as expressed by 1451 pupils in Bandung, a city of about one million inhabitants in West Java, to serve as a basis of comparison with opinions expressed by their own pupils.
2. To furnish educators and child psychologists in other countries data concerning this aspect of modern Indonesian children's development.

Setting of the Study

The pupils taking part in the study attended schools recommended by the district inspectorate as probably being rather typical of the Bandung school population (four elementary, four junior-high, and three college-preparatory senior high schools).

Although compulsory, universal education is an ideal of the new Republic of Indonesia (a nation made up of most of the islands formerly composing the Dutch East Indies colony), the ideal has not yet been realized because of a shortage of school buildings. An inspection of the occupations of the fathers of the pupils in the study suggests that the children in school are not representative of the general population of children. It appears that, especially in the junior and senior high schools, the sample is biased in favor of the upper-middle and upper socio-economic classes.

The total sample consisted of 831 boys and 620 girls. As seen in Tables I and II, the ratio of boys to girls in the elementary grades was about even, but in the junior-high and high schools there were between 35 percent and 110 percent more boys than girls at each grade level. This predominance of boys in the secondary school appears typical of schools throughout Indonesia.

Results of the Study

Substantially more likes than dislikes concerning school were expressed by both girls and boys. The mean number of likes cited by girls was 1.7, and by boys was 1.8. The mean number of dislikes cited by both grils and boys was 1.1.

The Most Popular Likes: The greatest percent of likes of both girls and boys related to school work, usually cited by subject-matter name. The following subjects were found among the first ten most popular choices of girls and boys alike: arithmetic and algebra, history, physical education, biology, English language, Indonesian language, and drawing. An attraction to the general activity of studying was also included among the first ten in both lists. (Tables I and II)

---

*All footnotes will appear at end of article.

TABLE I

THE TEN THINGS GIRLS LIKED MOST ABOUT SCHOOL*

Grade	3	4	5	6	7	8	9	10	11	12	Total N	Overall %
N	24	101	86	74	69	50	106	34	38	38	620	
Arithmetic and Algebra**	12. 5	12. 9	37. 2	37. 8	39. 1	35. 0	12. 0	5. 9	7. 9	5. 3	129	20. 7
History	—	7. 9	10. 5	28. 4	20. 3	12. 0	8. 5	23. 5	10. 5	15. 8	86	13. 9
Studying	33. 3	10. 9	17. 4	21. 6	20. 3	10. 0	12. 3	2. 9	—	—	83	13. 4
Physical Education	8. 3	2. 9	11. 6	4. 0	14. 5	28. 0	16. 0	8. 8	5. 3	2. 6	79	12. 7
Biology	4. 2	11. 9	16. 3	27. 0	11. 6	6. 0	10. 4	11. 8	5. 3	—	75	12. 1
Reading	12. 5	45. 5	16. 3	5. 4	—	—	—	—	—	--	67	10. 8
English language	—	—	—	—	23. 2	14. 0	16. 4	14. 7	29. 0	7. 9	59	9. 5
Indonesian language	—	—	2. 3	16. 2	13. 0	6. 0	17. 0	11. 8	15. 8	10. 5	58	9. 3
Music (singing and music theory)	8. 3	7. 9	14. 0	5. 4	13. 4	10. 0	4. 7	8. 8	—	--	55	8. 8
Drawing	12. 5	16. 8	12. 8	13. 5	2. 9	6. 0	. 9	2. 9	2. 6	5. 2	51	8. 2

* Figures in the table are reported as the percent of girls at each grade level who named that item as one of the things she liked most about school. The final column at the right shows the overall percent of girls, out of a total of 620, who named each item in the list. Following these first ten choices were 31 less popular ones, shown in the following list. (The number after the name of each choice is the total number of girls at all grade levels who chose this item as something liked). Geography (47); play with school friends (33); geometry (33); writing (25); physical science (25); exchanging ideas with others (23); paying attention to lessons (16); free period because the teacher did not arrive (16); handcrafts (15); Sundanese language - the local language of West Java (13); German language (12); business arithmetic (12); character education (6); friendly teachers (6); study of law (6); having a good recitation (5); hearing a story (5); chemistry (5); French language (5); buying candy at roadside eating stand near school (5); economics (3); civics (3); having a quiet class (3); morning classes (3); religious education (2); solving problems (2); bookkeeping (2); mechanics (2); writing compositions (1); health education (1); rote learning (1).

** Figures from grades three through six describe arithmetic choices, grades seven through twelve describe algebra choices.

TABLE III

THE EIGHT THINGS GIRLS DISLIKED MOST ABOUT SCHOOL*

Grade	3	4	5	6	7	8	9	10	11	12	Total N	Overall %
N	24	101	86	74	69	50	106	34	38	38	620	
Arithmetic and Algebra**	4.2	3.0	3.5	6.8	10.1	26.0	14.2	23.5	13.2	2.6	61	9.8
Playing naughty games	20.8	18.8	24.4	12.2	11.2	—	2.8	—	2.6	2.6	67	10.8
Quarreling	—	8.9	14.0	20.3	8.7	—	1.9	14.7	2.6	7.9	53	8.5
Drawing	8.3	13.9	4.6	10.8	13.0	—	7.5	2.9	13.2	2.6	52	8.4
Unruly class	—	4.0	4.6	18.9	7.2	4.0	10.4	17.6	10.5	2.6	51	8.2
Music - Singing and theory	—	8.9	15.1	14.8	5.8	4.0	.9	—	—	—	33	5.1
Physical Science	—	—	—	10.8	11.6	6.0	10.4	5.9	—	—	32	5.2
History	—	6.9	8.1	9.4	1.4	6.0	4.7	--	—	—	29	4.7

* *Figures in the table are reported as the percent of girls at each grade level who named that item as one of the things she disliked most about school. The final column at the right shows the overall percent of girls, out of a total of 620, who named each item in the list.* Following these first eight choices were 39 others, shown in the following list. (The number after the name of each choice is the total number of girls at all grade levels who chose this item as something not liked.) Poor recitation (27); geometry (25); no lesson because teacher did not arrive (21); geography (20); hearing a story (20); having the teacher angry with me (16); physical education (16); English language (15); Sundanese language (13); being bothered by companions (12); chemistry (10); playing with school friends (9); handcrafts (8); conceited people (7); naughty companions (6); Old Javanese language (6); business arithmetic (5); confusing lessons (5); buying candy at roadside stand (4) ; biology (4); reading (4); bookkeeping (3); economics (3); German language (3); Indonesian language (3); study of law (3); advice rom teacher (2); civics (2); children who copy others' work (2); dirty school (2); having to sit next to a boy (2); last class period of the day (2); tattle-tales (2); writing (2);afternoon classes (1), being a truant (1); enthnology (1); getting home late (1); mechanics (1); memorizing material in school (1); religious edu-ation (1).

* Figures from grades three through six describe arithmetic choices; grades one through twelve describe algebra choices.

TABLE IV

THE EIGHT THINGS BOYS DISLIKED MOST ABOUT SCHOOL*

Grade	3	4	5	6	7	8	9	10	11	12	Total N	Overall %
N	27	84	85	82	94	110	166	71	56	56	831	
Quarreling	14.8	31.0	34.1	13.4	11.7	8.2	6.0	9.9	—	8.9	112	13.5
Playing naughty games	14.8	29.8	28.2	18.3	18.1	6.3	9.0	—	1.8	—	108	13.0
Unruly class	—	7.1	11.8	11.0	7.4	11.8	5.4	19.7	5.4	8.9	76	9.1
Music - singing and theory	3.7	6.0	8.2	19.5	5.3	7.3	9.6	1.4	—	—	59	7.1
No lesson because teacher did not arrive	—	—	—	—	7.4	10.9	10.2	9.9	3.6	5.4	48	5.8
Drawing	11.1	2.4	4.7	9.8	3.2	1.8	6.6	4.2	3.6	1.8	39	4.7
Poor recitation	—	—	1.2	—	—	2.7	5.4	11.3	21.4	10.8	39	4.7
Arithmetic and Algebra**	3.7	—	2.4	3.7	—	8.2	4.8	2.8	17.9	—	35	4.2

* Figures in the table are reported as the percent of boys at each grade level who named that item as one of the things he disliked most about school. The final column at the right shows the overall percent of boys, out of a total of 831, who named each item in the list. Following these first eight choices were 46 others, shown in the following list. (The number after the name of each choice is the total number of boys at all grade levels who chose this item as something not liked.) Having the teacher angry with me (27); geometry (23); conceited people (22); English language (21); physical science (20); history (18); handcrafts (17); Sundanese language (14); German language (13); hearing a story (12); being bothered by companions (12); physical education (11); geography (11); biology (11); chemistry (10); playing with school friends (10); naughty companions (9); economics (8); business arithmetic (8); uninteresting teacher (7); astronomy (6); memorizing material in school (5); Indonesian language (5); having to sit next to a girl (5); tattle-tales (5); writing (5); children who copy others' work (4); buying candy at a roadside stand (4); dirty school (4); Old Javanese language (3); confusing lessons (3); reading (3); study of law (3); not passing to next grade (3); getting home late (2); losing school supplies (2); homework (2); rote memorization (2); studying (1); too expensive textbooks (1); ethnology (1); last class period of the day (1); wearing a uniform (1); having a friend sick (1); not having friends (1); religious education (1).

** Figures from grades three through six describe arithmetic choices; grades seven through twelve describe algebra choices.

they liked and disliked most about school. Although, pupils expressed a greater total number of likes than dislikes concerning school, there was a slightly greater variety of dislikes than likes expressed.

The likes focused primarily on school subjects, with arithmetic and algebra, history, physical education, biology, English language, Indonesian language, and drawing found among the ten most popular choices of both boys and girls. Girls showed a greater preference for reading, biology, music, and history, whereas a larger percent of boys than girls selected arithmetic and algebra, geometry, and physical education.

For both boys and girls the most common dislikes centered on disturbances in personal-social relations (quarreling, naughty games, unruly class) and certain school subjects (such as music, drawing, arithmetic, algebra, and geometry).

Two questions that arose during the analysis of the results bear mention:

1. Since in many ways the pupils' answers reflect the same kinds of likes and dislikes adults might hope children would have, is it not probable that the likes and dislikes expressed by these pupils indicate what they think their teachers prefer rather than what the students themselves really believe?

We cannot answer this question with certainty but we can note several factors that could have influenced the choices. First, although the pupils signed their names to the questionnaires and thus knew they could be identified, it was made clear to them by the administrator of the questionnaire (a research assistant unrelated to the pupils' schools) that the answers were for his use and would be seen by him alone, not by the teacher or school officials. Second, it should be recognized that children in Indonesian schools are likely to be ones that accept adults' standards, because schooling is a privilege and those who do not like school or do not do well in school are soon eliminated. Third, it should be noted that even though there was much agreement between children's expressed likes and dislikes and those that adults might wish them to express, the agreement was not at all perfect. Many, if not all, of the pupils seemed to feel free to express a dislike for regular school subjects in addition to expressing such adult-sanctioned dislikes as those relating to naughty games, an unruly class, and the teacher being absent.

Therefore, even though some pupils may not have expressed their real feelings, it appears logical that the general picture of likes and dislikes is probably fairly accurate. For children who have not succeeded in school and thus have not been allowed to continue, we might well expect a different picture.

2. Why are some of the most popular likes also the most common dislikes (such as arithmetic and algebra, geometry, drawing, music, and history)?

There are at least two factors probably operating to cause this. First, certain subjects (such as arithmetic and mathematics) are pursued by students throughout their school career, whereas other more specialized subjects (French language, economics) are taken only by a selected few students for a short time. Thus the subjects followed at several grade levels have more chance to be liked and disliked than do those met by only a small percent of the pupils. (This factor of opportunity to meet certain studies also helps account for some age trends: for example, English language is taught only at the secondary levels so it would not figure in the choices of elementary-school pupils.

However, this opportunity to meet a subject is not the only factor involved. For instance, children face reading and writing at all levels, but reading and writing did not figure prominently in likes and dislikes. Therefore, in addition to there being a difference in the frequency with which pupils encounter different subjects, there is apparently also a difference in the amount of emotion, either positive or negative, that certain subjects tend to generate. Because such activities as drawing, singing, and dealing with algebraic concepts seem to involve somewhat specialized abilities, it is possible that they cause great satisfaction for the successful pupils who possess these abilities and at the same time cause greater frustration for those who do not. It is possible also that children who are inadequate in more basic academic skills such as reading and writing have already been eliminated from school, whereas pupils with lesser abilities in music, art, and algebra can still follow the typical academic course successfully. More specialized research concerning pupils' abilities and teaching methods would be necessary to answer the questions involved here.

FOOTNOTES

* This study was carried out in 1959-1960 under a teacher-education project of the State University of New York, financed by the Ford Foundation. The aid of Research Assistants Suparman and Hutabarat is gratefully acknowledged.

1. Differences between percents significant beyond the one percent level of confidence.

2. Differences between percents significant between the one and five percent levels of confidence.

3. Differences between percents significant between the five and ten percent levels of confidence.

4. Differences between percents significant beyond the one percent level of confidence.

5. Differences between percents significant between the one and five percent levels of confidence.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 2, October 1962)

A Study of Morale in Education Utilizing Incomplete Sentences

JOHN H. SUEHR
Alamosa High School
Alamosa, Colorado

THIS STUDY IN EDUCATION parallels studies of morale which have been made in industry. Teacher morale in the Boulder, Colorado Public Schools was investigated, and correlated with personal adjustment of principals and satisfaction of parents. In addition, certain biological, sociological, and psychological factors were examined in relation to morale.

Design

Teacher morale. Investigators generally agree on the basic job factors contributing to morale; however, they differ concerning methods for measuring morale. Some maintain that low or high morale in the case of teachers is a product not only of environmental conditions, but also of the personality of the teacher. Thus, an instrument was needed that not only would reveal environmental factors affecting morale, but in addition would give an indication of the respondent's personality structure.

A study of the literature of morale psychology indicated a better method of measuring morale might be devised than the questionnaire type of instruments presently in use. An instrument of the incomplete sentences type was constructed for measuring morale. From the literature, from conferences with educators, and from personal experience, one hundred incomplete sentences were created as likely items to measure morale. These one hundred items were then submitted to 67 teachers from various parts of the country for choice of items most suitable. From their selection, forty items shown in Table I were chosen. This instrument, submitted to a pilot study, was found satisfactory for measuring morale.

Personal information. In order to gain from teachers information about psychological, sociological, and biological factors that might relate to morale, a questionnaire was devised. Almost every item utilized had been subject to prior research in areas other than teacher morale. Included were such areas as body-type, socio-economic level of the childhood home, and personality self-concepts.

Principal's personal adjustment. Because the emphasis in this study is on the incomplete sentences method, the Incomplete Sentences Blank — referred to as the ISB — was selected to determine personal adjustment of principals. This test was devised by Julian Rotter and Janet Rafferty and published through the American Psychological Corporation.

Parent opinion. One would naturally think that teacher morale would have an effect on parent opinion, but prior to this study no researcher had attempted to measure this relationship. The writer designed a double postcard that simply asked parents to state the degree to which they were satisfied with the job the school was doing for their child. A five percent sample was selected as sufficient for Boulder.

Presentation and collection. The morale study was scheduled for the same day in all the Boulder schools. Prior to this day, meetings were held with the principals and superintendent. The teachers selected one of their group to collect the forms and mail them to the School of Education at the University of Colorado. While the teachers were completing their forms, each principal was taking the ISB.

The reaction of the teachers to the Incomplete Sentences Form was good; one teacher remarked on the cathartic effect. Teachers often are plagued by requests for participation in doctoral research studies. In spite of probable satiety, the teachers expressed their appreciation in being allowed to express and formulate their own thoughts.

The enrollment cards in each building furnished the names of parents. Fisher's Tables of Random Numbers were used to draw a porportionate sample stratified on basis of sex. Each principal was shown the names of the parents selected, and in every case they felt that a random sample had been drawn. As a result of the first card, 60 percent of the parents responded. A follow-up card was sent two weeks after and an additional 14 percent responded. Those who responded after the second request showed a greater degree of dissatisfaction.

Scoring. In arriving at a scoring procedure for the Incomplete Sentences Form for measuring teacher morale, an arbitrary decision was made to have five scoring categories ranging from high to low morale, or from zero to four. Three scorers were selected on the basis of training, public school experiences, general perceptiveness, and recommendations of the advisory committee. There was no contact between scorers; the only instructions given to each scorer were these:

Instructions for Scoring the Incomplete Sentences Form

1. Read each statement over carefully.
2. Assign each statement a point value based on the scale below.

   0 - Highly positive statements — denoting high morale.
   1 - Slightly positive statements.
   2 - Neutral statements.
   3 - Slightly negative statements.
   4 - Highly negative statements — denoting low morale.

3. If a statement is both positive and negative, weigh the two, and decide if one is stronger than the other. If one does not predominate, score the statement as neutral.
4. Total score for each form will be computed by machine.

The interscorer reliabilities are presented in Table II. While there can be no doubt that the scorers in this case were highly trained people, personnel with school experience should score forms with consistency.

An example of typical completions and how they were scored follows.

20. American education . . .

(4) reflects our ease and love of tranquilizers; is bogged down by too many kids who do not desire it; should improve in all areas if we are to exist as a nation

(3) needs more emphasis on humanities; is trying to do too much; stresses efficiency, not character; is excellent in concept but below in practice; needs something

(2) is neither as bad as Mr. Bestor would hope nor as good as the NEA feels; has, after all, the masses to consider; should never be compared with European

(1) on the whole is fine but it must get better; has done much but has a long way to go; is good for the intelligent; is trying to improve

(0) is the best for the most and that's as it should be; is generally good; is the best in the world; serves the purpose for which it is intended; is a wonderful system

An item analysis of the Incomplete Sentences Form for measuring teacher morale showed that only six items failed to discriminate between high and low morale teachers at the one percent level of significance. One of the forty items, "My class size," failed to discriminate at the five percent level. This clearly demonstrates the discriminative power of the incomplete sentences method as an instrument for assessing teacher morale.

Since all the parent opinion cards were coded as to school, it was relatively simple to group the cards by buildings. In keeping with the plan used in scoring the teacher morale forms, "very satisfied" parents were scored as zero, "satisfied" as one, "no opinion" as two, and "dissatisfied" as three. By assigning each card a score, adding the total, and computing a mean, the schools in the system were ranked on parent satisfaction.

The Validity Problem

Before studies of morale in education can compare with those of industry, validity of measures must be established. The question of morale in education is complicated by the lack of objective criteria. While no specific validity system applies perfectly to this problem, construct validity is the most pertinent. For purposes of dem-

TABLE I

THE ITEMS AND INSTRUCTIONS FOR THE INCOMPLETE SENTENCES FORM FOR MEASURING TEACHER MORALE WITH INSTRUCTIONS

Complete the following in order to make each one a complete sentence. Take your time and show your true feelings about your situation. All responses will be confidential.

1. Teaching school . . .
2. My salary . . .
3. My future in teaching . . .
4. Morale of teachers . . .
5. Faculty meetings . . .
6. The people of this community . . .
7. My working environment . . .
8. Teachers . . .
9. My principal . . .
10. Teamwork among teachers . . .
11. Children today . . .
12. Parents . . .
13. Fringe benefits in teaching . . .
14. Channels of communication . . .
15. Clerical help . . .
16. The school board . . .
17. Administrative policies . . .
18. Professional organizations . . .
19. Evaluation of my work . . .
20. American education . . .
21. Opportunity to help make policy . . .
22. The PTA . . .
23. Teaching materials . .
24. Custodians . . .
25. Professional standards . . .
26. My pupils . . .
27. This community . . .
28. The superintendent . . .
29. Discipline . . .
30. Teaching assignments . . .
31. Personnel policies . . .
32. School administrators . . .
33. This school . . .
34. My class size . . .
35. The in-service program . . .
36. Teacher welfare . . .
37. The curriculum . . .
38. Teacher opinions . . .
39. My teaching ability . . .
40. Teaching again . . .

TABLE II

INTERSCORER RELIABILITY ON THE INCOMPLETE SENTENCES FORM FOR MEASURING TEACHER MORALE BETWEEN SCORERS A AND B, A AND C, B AND C

Scorers	Pearson r
A and B	.97
A and C	.99
B and C	.98

onstrating the validity of the Incomplete Sentences Form, the following are offered as indications of construct validity.

1. Evidence is given of maladjustment among the low-morale teachers.
2. Evidence of internal consistency is shown by the interscorer reliability figures in Table II.
3. Low-morale teachers did miss more school than high-morale teachers. This is often referred to as validating morale indicators used in industry.
4. An administrator familiar with all the schools in the system predicted schools in the extreme areas of morale.
5. An item analysis of the Incomplete Sentences Form demonstrated its aptitude in discriminating between teachers.
6. Item No. 4 of the ISF gives a picture of how well teachers can tell if morale is high or low in a specific building. The coefficient of correlation between ranking schools by the ISF and composite teacher opinion measured by Item No. 4 was .53.

Analysis of Data

The data collected were analyzed by the customary statistical method. Chi-square was used to test for significant differences between the upper and lower quartiles on the morale scale. Because an extremely large number of tables would be needed to present the data adequately, only the findings are presented here.

Findings

Parents. Parent satisfaction was significantly related to teacher morale. This relationship could not be attributed to level of education or to socio-economic factors.

Principals. Personal adjustment of principals did not relate significantly to teacher morale. It is to be noted that no principal furnished responses placing him in the poorly adjusted category. Of course, as psychologists have said, the greatest limitation of personal adjustment as a norm is that it omits any reference to the individual's role in his group. In some cases, the principal with a good score for personal adjustment was not well thought of by his teachers.

Teachers. These factors were found to relate significantly to morale at the five percent level.

1. More males than females were found in the low-morale group, in the ratio of three to one.
2. Teachers indicating that an opposite-sexed parent had the most influence on them were more apt to be low-morale teachers.
3. Knowing or estimating IQ to be high would increase probability of low morale.
4. Being the youngest in the childhood family was a factor in low morale.
5. More days of school were missed by low-morale teachers.
6. High-morale teachers feel that they are realizing their fullest potential in teaching; low-morale teachers do not.
7. Low-morale teachers consider themselves either more or less gregarious than do high-morale teachers.
8. Teachers of the high-morale group have taught longer than their opposites.
9. High-morale teachers worry less than do low-morale teachers.
10. High-morale teachers feel that in teaching they have fulfilled their parents' expectations for them. Low-morale teachers do not.
11. Low-morale teachers feel their personal appearance is above average; high-morale teachers more often indicate average personal appearance.
12. Socio-economic status of one's childhood family plays an important role in future morale in teaching. High-morale teachers come from upper or upper middle class homes more often than low-morale teachers.
13. Low-morale teachers feel their personality leans more toward one of superiority in its outward manifestation.
14. Low-morale teachers feel they are more stubborn in their personality make-up.
15. High-morale teachers more often indicated that both of their parents were happy in their respective occupations.

16. High-morale teachers are much more apt to have grown up in an urban society; rural origin was three times as common in the low-morale group.

17. High-morale teachers more often feel that their childhood family was very close.

18. High-morale teachers are more apt to go to bed early and to get up early.

19. High-morale teachers feel that they have more close friends.

20. Low-morale teachers consider their self-confidence greater than average, while high-morale teachers more often consider it average.

21. Low-morale teachers feel that they repress their true feelings less than the average person.

22. Low-morale teachers more often feel above or below average in their degree of perseverance.

23. Low-morale teachers rate their degree of ambition as significantly greater than average.

24. Low-morale teachers more often rated their use of alcoholic beverages as being more than average or less than average.

25. High-morale teachers were more apt to rate themselves as being slightly introverted, and low-morale teachers were more often found in the slightly extroverted category.

26. High-morale teachers were more often stout or plump; low-morale teachers indicated a muscular body-type.

Conclusions

The conclusions reached by the writer as a result of this study are:

1. The childhood background of teachers is highly important to their morale as adult teachers. High-morale teachers have an advantage from the beginning by more often being born into higher socio-economic status. Research in sociology has previously demonstrated this point.

2. As among students, personality differences exist among teachers.

3. Cause-and-effect relationships are evident in the area of morale. Parents sending their children to schools in which teacher morale is low show a corresponding doubt of that school.

4. The incomplete sentences method seems to be the best method yet devised for determining teacher morale level. It uncovers a broader area of "trouble spots," and is more discriminating.

5. Communication is probably the most vital area in the whole morale process.

6. Much is already known about pupil motivation, but little is known about what motivates teachers. The problem is important.

7. Human relations seem to be the key concept to morale.

8. Teaching is not adequately recognized as a desirable profession for a man; consequently, the type of man entering public school teaching might possibly be different from the man entering medicine, law, or business.

9. Low or high teacher morale is the result of a holistic situation. The determinants of a teacher's attitude are pluralistic and interactional, and all possible relevant factors — biological, psychological, and sociological — must be investigated and coordinated into a better understanding.

10. A conclusion not unique to this study is the fact that the mental health of the teacher is important.

Recommendations

For teacher training. The results of this study could be used in a structured interview with all prospective teaching candidates. A student fitting into a pattern leading to low morale in teaching should either be counselled out, or given, through therapy, insight into his conflicts.

Prospective teachers should have a trial in teaching earlier in their college career. Educators are often so eager to recruit teaching candidates that quality of personality for teaching is overlooked. Teacher training institutions should be selective on the basis of many factors.

For teacher selection and placement. A highly structured interview should be an integral part of the selection procedure. Interviews could be more valuable in teacher selection than they are now.

Industry is presently experimenting with the use of personality tests for determining what types of personalities work best with each other. Results seem to be successful. Education should heed this consideration; use might be made of findings in industry. Some teachers, for example, want to be democratically administered while others are more secure under an authoritarian type of administration. In pupil placement, education already does this on an unscientific basis. Certain pupil personalities are often given to selected teachers.

For teacher rehabilitation. This year, for the first time, the Boulder Public Schools are partially paying for physical examinations for teachers. Importance is attached to physical health; the time should come soon in which mental health is considered of equal importance. Teachers obviously needing help should be provided with psychiatric aid. This would involve cost, but the returns would more than make up for it. The possibility of group therapy sessions should be explored for their savings in cost.

For parent satisfaction. The PTA lowered teacher morale more than any other factor. The PTA should be thoroughly studied. Perhaps a better method of bringing home and school together could be instigated.

For administrators. Every administrator, either as part of his training, or as in-service training, should undergo activities aimed at greater self-understanding. One must first have self-understanding, before he can understand and be interested in others and their problems.

For education generally. 1. A procedure for dealing with grievances should be operating in every school system. Whether the grievance procedure is often used or not is of little concern; what matters is that it is there.

2. All schools should make periodic morale check-ups.

3. There is too much haphazard encouragement of people to become teachers in America. Classrooms will not go empty very long. The people will raise salaries and the prestige of teaching will come up, and with it, teacher morale. Raising standards — not those of IQ but those of quality of the individual — will raise the profession and attract able people. Coordinated action by all teacher training institutions will be needed.

4. In most organizations there is an intolerance of internal criticism. In this study and other morale studies, lack of communication on all levels was evident. Institutes of industrial relations at various universities throughout the country are beginning to propose a type of activity where personnel within an organization sit down and freely discuss each other and their problems. For the period of the session, top management and workers are on equal terms. The aim of this training is to enable one to see what kind of person others think he is, and what impact he has on others. A recommendation of this study is the application of this procedure in the public schools.

Recommendations for Further Research

Incomplete sentences method. For greater discriminative power and for richness of material, the incomplete sentences method should be used in more research studies. The breadth of material obtained more than compensates for the trouble of scoring. Instruments for measuring student morale, administrator morale, parent satisfaction, and feelings on attitudes of non-certificated personnel should be devised.

Morale check-ups. Periodic measurement of morale should be part of every school-year schedule. Scoring could be done by teacher judges elected by the staff. A manual to facilitate scoring for the Incomplete Sentences Form used in this study will be published.

Stability of morale. A very worth-while, though difficult-to-structure study, would be to determine whether or not morale varies during the school year. The writer feels that the results of the Incomplete Sentences Form will not vary to any great extent during the year.

Production and morale. Since the main feasible way presently known in education to determine "production" is through the results of achievement tests, a study should seek to find out whether or not students in high-morale schools are achieving in consonance with their ability.

Action research. Action research in this area is badly needed. Studies to determine what techniques will raise morale should be made. The principal of a high-morale school might trade positions with one of a low-morale school and the effect be noted. A school of low morale should be subjected to the awareness-type activity previously mentioned. A school having low morale should have its personnel undergo therapy to see whether or not better self-understanding raises morale. All personnel of a low-morale school should undergo personality testing, and replacements or transfers made on the result of these tests; in essence, people with personalities that harmonize should be assigned to a low-morale school and the results evaluated.

In any experiment, the effect not only on teacher morale, but also on pupil, administrator, and parent morale as well, should be evaluated.

More thorough investigation. The findings of this study are largely suggestive in nature. While it is true that this type of study is needed at this time, each of the findings should now be subjected to intensive, direct research. Only the early studies should be broad in scope. There is no subject more worthy of investigation in education than morale, but it will take far-sighted, imaginative people to accept the challenge adequately to study this area.

REFERENCES

1. American Association of School Administrators. Morale for a Free World. Twenty-second Yearbook (Washington: The Association, 1946).
2. Anderson, Harold H., and Anderson, Gladys L. (ed.). An Introduction to Projective Techniques (Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1951).
3. Griffiths, Daniel E. Human Relations in School Administration (New York: Appleton-Century-Crofts, Inc., 1956).
4. Roethlisberger, F. J. Management and Morale (Cambridge, Massachusetts: Harvard University Press, 1941).
5. Rohde, Amanda R. The Sentence Completion Method (New York: The Ronald Press Company, 1957).
6. Rotter, Julian B., and Rafferty, Janet E. The Rotter Incomplete Sentences Blank (New York: The Psychological Corporation, 1950).
7. Weschler, Irving R., et al. "Yardstick for Human Relations Training," Adult Education, VII, No. 3 (Spring 1957), pp. 152-168.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 2, October 1962)

Teacher Time Utilization With Teacher Aides

CURT STAFFORD
San Jose State College

EDUCATION IS EXPERIENCING a rumbling deep in its interior. It may be on the verge of a qualitative change in some of the deeply seated concepts of teacher/learner roles and of scheduling the program of studies. One of the more provoking suggestions for these changes is Trump's *Images of the Future* (1). An element of this forward look is the use of clerical aides to allow the teacher to be concerned with more professional matters. Since education requires that ideas be supported by research, there is an obvious need for study of the clerical aide.

The use of lay persons to free the teacher of clerical tasks was a major aspect of the Cooperative Study for the Better Utilization of Teacher Competencies. Sometimes known as the "Bay City Project," this program was carried on at Central Michigan University under a Ford Foundation grant. Journal articles concerning the teacher aide program have been authored largely by persons external to the program who were invited to observe it and report their considered opinions (2). The majority of reports by the staff itself consisted of locally printed materials (3), although a description of it did appear (4). The present writer was a member of that staff and feels strongly that a report to the profession of smaller segments of the study is long overdue.

The teacher aide program was initiated as a partial solution to the problem of increased enrollment in the face of the teacher shortage and inadequate classroom facilities. Most of the controversy over the program centered on the lowering of standards of teacher certification and so insufficient attention was paid to the program's potential contribution to a fundamental reorganization of the role of the teacher and his classroom. This article has two purposes: (1) report the effect of the teacher aide on the elementary teacher's time utilization, and (2) point up the need for deeper consideration of the definition of teacher and aide roles as professional educators attempt application of plans such as Trump's. The relationship of teacher and aide is important irrespective of the grade level at which it exists.

Method of the Time Study

Two field staff members performed a job analysis of the elementary school teacher. This analysis resulted in the definition of twenty-one in-school activities. They are:

Lesson Plans: Time used by teachers in the preparation of lessons plans and the reference to plan books, and in the making of notes from text material. Preparation of master copies for duplicator use is included in this category.

Assignments: Time used in presenting advance lessons.

Group Planning: Time used in the discussion of the future lessons or in the evaluation of work in progress or completed. (Pupil participation is the sole differentiating factor between this category and "Assignments.")

Counseling: Time in which the teacher is engaged in individual or group counseling at the opening or closing of the schools session. (Pupil control is not included here.)

Write on Board: Time spent in the mechanics of putting material on the board. This is done, in most instances, before or after the school session.

Written Lesson: Time spent in monitoring pupil activity, including written work and examination.

Report Cards: Time spent in clerical activity such as recording grades or other data on reports to parents. It does not include time spent in evaluating progress or pupil growth and development.

Recitation: Time during which the teacher is engaged in discussion with a pupil group, including read-back of lesson materials, free discussion and experience-sharing by pupils.

Directed Study: Time used in assisting pupils who are engaged in group study.

Desk to Desk: Time in which the teacher goes from desk to desk giving individual help.

Correct Papers: Time spent in the clerical aspect of checking or correcting papers, written lessons, workbooks, notebooks, and tests.

Pupil Control: Time given to establishing disciplinary control over an individual or a group.

Taking Roll: Time used to record absences and tardiness.

Opening: Time used for prayers, the pledge of allegiance, group singing of morning songs, and other routine announcements.

Interruptions: Time involved in such activities as messengers from other rooms, visits by other teachers, administrative personnel and special education workers, parent interviews, telephone messages, collection of funds, and fire and air raid drills.

Housekeeping: Time spent in arranging for the physical comfort of pupils, including such items as the hanging of pictures, and the placing of posters and seasonal decorations, the care of blackboards, bulletin boards, display cases, and the securing, distributing and collection of papers, materials and tests.

Making Reports: Time devoted to making school reports such as attendance reports, anecdotal reports, entering data on the pupil's permanent record card, recording test results, and the making of requisitions and inventories.

Dictation: Time spent in dictating spelling lists and similar materials to pupils.

Reading to Group: Time spent in reading instructional and recreational material to pupils.

Transition: Time consumed by the teacher in monitoring pupils as they go from seats to class and back or from room to room, or as they re-group. Included are pupil line-ups at drinking fountains, toilets, and lunchroom.

Supervision: Time spent in monitoring rest periods, free play, recess and halls. Time consumed when the teacher remains with her group and gives attention to the scheduled activities of special education personnel such as art, music, and physical education (5).

A time study of the teacher's day was made by two field staff members who accompanied teachers throughout the entire day. The staff member noted what a teacher was doing and classified this according to the list of defined activities. The duration of the activity was then measured with a stopwatch calibrated in hundredths of a minute. "Time utilization" is the profile of time devoted daily to each of the twenty-one defined activities.

Reliability

The time study was performed in order to obtain an indication of total time utilization in elementary classrooms over a full school year. The accuracy of this estimate is indicated by the standard error of the mean. Table I reveals that error in reading the stopwatch certainly was insignificant. Furthermore, when a time sheet was reclassified after a one week interval, the proportion of the total day identically classified was in the vicinity of 98 or 99 percent. No significant differences between timers were discovered when their mean time utilizations were compared by the method of chi-square, nor did split half comparisons for timers reveal any significant variations within timers. Thus, it may be concluded that the measures are reliable.

Validity

Validity is always validity for a purpose and consequently the validity of the time study data becomes a function of how one attempts to use it. For the purpose of describing the time utilization of the elementary teacher, the time study data is valid because it is relevant. The definitions of the activities describe teacher actions observable from an external reference and make no reference to pupil learning. Distinguishing one activity from another is thus a relatively simple matter. For example, taking roll, write on board, and dictation are easily differentiated. The differences between some other activities may be more difficult to judge, as between assignments and group planning. On the whole, however, it is no particular task to classify in accordance with the definitions previously presented.

The data of the time study also have been considered in relation to the competence or professional training required to perform the activity. Validity for this purpose necessarily is less because activity definitions were not made to take into account the professional training required. Nevertheless, it is possible to make a reasonable estimate of the portion of the teacher's day devoted to activities either incongruous with the requirements of professional training or not justifiable in terms of the high cost of teacher time.

There are, then, two things which the time study data validly measure:

(1) average teacher time utilization in the activities defined, and
(2) (with somewhat less validity) proportion of time spent on activities which do not require extended professional training.

The time study data have little validity in the realms of teacher effectiveness and pupil learning because the definitions of the in-school activities carefully avoid reference to teacher effectiveness and thus lack relevance. This omission was considered necessary to gain acceptance among teachers who were to be objects of time study. Since there is no rating of effectiveness, when more time is available for planning, the teacher who plans well has more time to plan well — the teacher who plans poorly has more time to plan poorly. With the absence of evaluative aspects from the activity definitions it becomes necessary to assume that an increase in time available results in improving planning, an assumption which probably cannot be satisfied over a full range of teachers. If the data are to be considered as related to pupil learning, it is necessary to make the assumption that gain in time for activities considered to be vital to a good education program results in improvement of the program. Again the assumption cannot be satisfied over a full range of teachers.

Procedures

A time study was performed on 80 teachers of single graded classrooms in Bay City, Michigan, prior to the use of teacher aides. The averages for these 80 teachers provided a measure of "without-aide" time utilization. The mean classroom enrollment was 33. Time studies were performed on 18 teachers of single graded classrooms with aides during the next two years. Two teachers were timed for five consecutive days in 1956. This latter sampling cannot be regarded as random, but it is not particularly biased in view of the day to day variation in performance patterns of individual teachers. The averages of these 18 cases provided "with-aide" measures. The mean classroom enrollment was 48. The differences between the with-aide and without-aide time utilizations were examined by t-tests.

Aide time utilization also was obtained by timing 2 aides for 5 days each. Following a comparison of teacher time without and with aide, the aide time utilization profile was added to the aided teacher's profile under the assumption of total benefit to the child. Time per day per child was then computed for certain activities.

Results

The pre-aide and post-aide measures of time utilization, and the results of the t-tests of significance of difference, are presented in Table I.

Activities Requiring Professional Training

Activities which are considered to be closely related to instruction or which are in the domain of professional skills are: lesson planning, assignments, group planning, counseling, recitation, directed study, desk-to-desk, pupil control, reading, and supervision. If the aide is making the contribution expected, the time devoted to these activities by the teacher should increase. (Except, it would be hoped, pupil control.) Refer to Table IIa for the summary.

All three activities showing a significant increase appeared in this grouping. Most other activities gained as expected. The apparent increase of total time devoted by the teacher to activities requiring professional training was, however, not statistically significant.

Activities Incongruous with Professional Training

Activities which are clerical or monitorial are: write on board, written lesson, correct papers, taking roll, report cards, opening, interruptions, housekeeping, making reports, dictation, and transition. If the aide is making the contribution expected, the time devoted to these clerical and monitorial activities by the teacher should decrease. Refer to Table IIb for the summary.

Only two activities (write on board and report cards) failed to differ at least in the direction expected. Six activities showing significant decrease in time for this group of activities can be accounted for by chance less than one time in 30. Thus it may be said that the time spent on activities by the aided teacher which are incongruous with professional training is significantly decreased.

Aide Time Utilization

Table III summarizes the time utilization of the aide. The profile is about what was antici-

TABLE I

DIFFERENCES IN TEACHER TIME UTILIZATION: WITHOUT-AIDE AND WITH-AIDE TIME IN MINUTES DAILY

Activity	Without-aide Mean	Without-aide $\sigma\overline{X}$	With-aide Mean	With-aide $\sigma\overline{X}$	Mean Difference	σDifference	t
Lesson Plans	9.5	1.8	17.5	3.8	8.0	4.2	1.9*
Assignments	5.6	0.5	7.8	2.3	2.2	2.4	0.9
Group Planning	4.8	0.7	10.0	2.5	5.2	2.6	2.0*
Counseling	5.2	0.7	4.2	0.9	-1.0	1.1	0.9
Write on Board	3.2	0.5	3.2	1.3	0.0	1.4	0.0
Written Lesson	13.5	1.9	4.6	1.7	-8.9	2.6	3.4**
Report Cards	0.0	0.0	0.0	0.0	0.0	0.0	—
Recitation	104.3	5.7	137.1	7.8	31.5	9.6	3.3*
Directed Study	64.3	5.0	49.8	9.9	-14.5	11.1	1.3
Desk to Desk	22.9	2.6	24.0	4.2	-1.1	4.9	0.2
Correct Papers	11.1	1.6	4.0	1.5	-7.1	2.2	3.2**
Pupil Control	2.7	0.5	1.8	0.4	-0.9	0.6	1.5
Take Roll	2.4	0.2	0.4	0.2	-2.0	0.3	6.7**
Opening	5.8	0.6	2.7	0.5	-3.1	0.8	3.9**
Interruptions	9.8	1.2	9.4	1.9	-0.4	2.2	0.2
Housekeeping	11.0	1.2	10.3	2.0	-0.7	2.3	0.3
Making Reports	4.6	1.4	0.0	0.0	-4.6	1.4	3.3*
Dictation	4.4	1.4	4.1	2.0	-0.3	2.4	0.1
Read to Group	9.9	1.7	9.7	2.0	-0.2	2.6	0.1
Transition	23.2	1.3	12.3	2.0	-10.9	2.4	4.5*
Supervision	27.5	2.4	27.9	3.5	0.4	4.2	0.1
Free Time	14.1	2.1	18.7	2.4	4.6	3.2	1.4

* Statistically significant increase

** Statistically significant decrease

TABLE IIa

CHANGE IN TIME UTILIZATION FOR ACTIVITIES REQUIRING PROFESSIONAL TRAINING (PUPIL CONTROL NOT INCLUDED)

	Without-aide (N=80)	With-aide (N=18)	Difference
Mean Daily Time	254.0	288.0	34.0
σ $\overline{X}$	30.0	32.6	
σ Difference			44.3
t			0.8

Interpretation: In direction expected, but not a significant overall increase.

TABLE IIb

CHANGE IN TIME UTILIZATION FOR ACTIVITIES INCONGRUOUS WITH PROFESSIONAL TRAINING

	Without-aide (N=80)	With-aide (N=18)	Difference
Mean Daily Time	89.0	51.0	36.4
σ $\overline{X}$	9.9	12.4	
σ Difference			15.8
t			2.3

Interpretation: Significant Decrease

TABLE III

AIDE TIME UTILIZATION

(TIME IN MINUTES)

Activity	Mean	σ X̄	Activity	Mean	σ X̄
Lesson Plans	7.8	3.2	Pupil Control	0.1	0.1
Assignments	0.0	0.0	Taking Roll	4.9	0.7
Group Planning	0.0	0.0	Opening	1.4	0.4
Counseling	0.7	0.4	Interruptions	6.6	1.6
Write on Board	0.7	0.6	Housekeeping	90.4	11.4
Written Lessons	0.0	0.0	Making Reports	12.6	3.0
Report Cards	7.2	4.6	Dictation	1.0	1.0
Recitation	0.0	0.0	Read to Group	5.8	2.8
Directed Study	61.2	8.0	Transition	3.3	1.1
Desk to Desk	16.4	3.8	Supervision	51.4	0.5
Correct Papers	81.0	12.9	Free Time	7.2	0.2

TABLE IV

ACTIVITY TIME PER CHILD, WITHOUT- AND WITH-AIDE

(TIME IN MINUTES PER MEAN ENROLLMENT)

Activity	Without-aide (Mean Enrollment = 33)	With-aide (Mean Enrollment = 48)		
	Teacher*	Teacher*	Aide	Total
Desk to Desk	0.69	0.50	0.34	0.84
Directed Study	1.93	1.04	1.27	2.31

* The difference between teacher times is significant for the activity Directed Study (t = 3.4) but not for Desk to Desk (t = 1.6).

pated. Housekeeping and correcting papers account for almost half her time. Supervision and directed study account for another third.

Time Per Child

The quantitative measure of activity per child suggests that some qualitative changes in the teacher may have taken place with the advent of the aide. While the per child measure has definite limitations, it does seem to have some relevance in the matter of individual attention. Since individual attention may be regarded as occupying the larger part of the activities desk to desk and directed study, the per child figures for these activities are presented in Table IV.

These data suggest that the teacher-with-aide is not able alone to provide a comparable amount of individual attention. However, if the kind of attention the teacher alone provides in an aided situation is qualitatively different, it may not be important to maintain the same time per child. For instance, perhaps these activities have taken on the character of pupil self-direction or small group work. The time study data do not measure along this qualitative continuum, however, so this contention remains unsupported.

A more plausible explanation is that the teacher has attempted to maintain the same ratio (6) and kind of individual attention by directing the aide to share in this activity. Thus the teacher's role is altered in that he is not the sole provider of individual attention. Another kind of qualitative change may be present, one in which the teacher now ministers professional skills through the medium of the aide. Once again, however, the time study data do not tell us if these changes are taking place.

The employment of the clerical aide, then, is far from a simple and discrete reassignment of tasks. The aides in the situations studied clearly are integral parts of the education of the children, not mere appendages. If appears as if the teacher must share the class with the clerical aide. The present writer is not overly distrubed about this for it will require that only the truly self— understanding teacher be assigned an aide. In the long run this must result in increased professional standards.

Summary

Elementary teacher time utilization was compared with utilization under conditions of increased enrollment and the measures of time, teacher activities requiring professional preparation failed to show a significant increase. Activities of a clerical nature, however, decreased significantly.

The teacher alone was unable to maintain individual attention in an amount comparable to the pre-aide situation. Individual attention in at least the same ratio could be provided if the aide is permitted to share in this activity.

This study was concerned solely with the quantitative aspects of time utilization — there is need for an investigation of possible qualitative changes.

FOOTNOTES

1. Trump, J. Lloyd. Images of the Future. Commission on the Experimental Study of the Utilization of the Staff in the Secondary School, National Association of Secondary School Principals.
2. See the June, 1956, issue of the Journal of Teacher Education for a collection of articles by Charles Park, Lucille Carroll, Francis Chase, James Hymes, Jr., Dorothy McCluskey, G.E. Rast, and P.G. Rulon.
3. A Cooperative Study for the Better Utilization of Teacher Competencies. Department of Special Studies, Central Michigan College. First Printed Report, November, 1953. Second Printed Report, 1955. Third Report, June, 1957.
4. "The Teacher Aide Plan," The Nation's Schools, July, 1955, pages 45-55.
5. First Printed Report, pages 13-14.
6. The figures totaling the teacher and the aide in Table IV appear larger than those for the teacher alone, but the differences are not statistically significant.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 2, October 1962)

Factors Involved in the Adjustment of Children Entering First Grade

SARAH LOU HAMMOND

DORA S. SKIPPER

Florida State University

THIS EXPLORATORY STUDY is an attempt to identify factors involved in the adjustment of the child in the first-grade classroom. First-grade teachers evaluated adjustment of 1,053 children. Twenty-five percent of the children having the highest and 25 percent having the lowest adjustment scores were selected for study.

The study is specifically concerned with how children of high and low adjustment compare in regard to such factors as: 1) attendance at kindergarten, 2) scores on readiness tests, 3) ordinal position of the child in the family, 4) size of family, 5) socio-economic status, and/or 6) chronological age of the child.

Procedure

Experimental situation. The subjects for this study were selected from the first grades of eight elementary schools in Orange County, Florida. These schools were selected on the basis of willingness to cooperate.* Each of the 32 first-grade teachers in these eight schools completed a questionnaire for each child in her classroom. A total of 1,053 questionnaires were completed by the teachers and scored by the investigators.

From this group of 1,053, the 25 percent of the children having the highest adjustment scores, a total of 260, and the 25 percent having the lowest adjustment scores, a total of 260, were selected. For each child in the two groups, information relating to age, position in the family, occupation of father and mother, education of father and mother, attendance in kindergarten, and scores on the Metropolitan Readiness Tests was collected. Complete information was obtained for 234 children having high-adjustment scores and 221 children having low-adjustment scores.

Sources of Data. The instrument developed by Lyle, "Questionnaire to Determine Feelings of First Grade Teachers Concerning Kindergarten Experiences of Children," (10) was adapted for use in this study. The instrument for this study consisted of five pages of mimeographed statements grouped according to the three areas of social and emotional adjustment, physical development and health, and readiness for learning. Teachers were asked to indicate for each child whether or not they observed *No*, *Some*, or *Much* evidence of development in each area. These forms were completed by the teacher at the end of the first six weeks of school and scored by the investigators.

Teacher ratings of *No*, *Some* or *Much* for the questions in all three areas were totalled. Percentages were calculated for the *No* and *Much* ratings on each of the 1,053 questionnaires. The 260 children (25 percent) who received the largest number of ratings in the *Much* classification, were selected as having high adjustment to school, and the 260 (25 percent) receiving the largest number of ratings in the *No* classification were selected as having low adjustment to school.

The Metropolitan Readiness Tests, Forms S (7), were administered by the teachers during the first six weeks of school. Additional data regarding the child's age and position in the family, occupation and education of father and mother, and number of children in the family were secured by the teachers from the school records.

Socio-economic status of the family was determined from the above information by use of the McGuire-White Scale (11), involving the occupation, source of income, and educational attainment of the father, except in cases where the mother provided family support.

---

* The cooperation of Miss Pauline McDonald, Elementary Supervisor of the Schools in Orange County, Florida is acknowledged. Miss Lillian Zachary served as the principal research assistant.

Information was tabulated for the high-adjustment and low-adjustment groups and presented in tables which indicate both the percentages and chi-square values. Percentages were used to indicate the proportion of children in each category. Chi-square values were calculated to examine the relationship between adjustment status and the age of school entrance, attendance at kindergarten, scores on the readiness tests, socio-economic rating, family size, and position in the family for the appropriate degrees of freedom at the .01 level of confidence.

Limitations. 1. Thirty-two first-grade teachers participated in the study; the high and low adjustment scores were based on teaching rating. The initial rating was made for all children in the classroom without reference to age, kindergarten attendance, or socio-economic status.

2. In using the McGuire-White Scale, precise information concerning source of income was required. Since this information was not always available from school records, it was assumed that the income came from occupational earnings only.

3. No distinction was made concerning length of time the children attended kindergarten. The few cases that indicated the time spent in the kindergarten as less than a school year were included. Also included were 12 children who attended summer kindergarten sessions.

Results

Attendance at Kindergarten. From the information presented in Table I, it will be noted that 53 percent of the children having high adjustment status had attended kindergarten. Only 25 percent of the children having low adjustment status had kindergarten experience. Since public kindergartens have not been available in the county, except for one summer on a tuition basis, the children attended either private or church-related kindergartens.

The obtained chi-square value reveals that for those children there is a significant relationship at the .01 level between adjustment status and attendance at kindergarten. This indicates there is a relationship between high adjustment in the first grade and attendance at kindergarten, and the first question is answered affirmatively.

These data support findings of Morrison (12). In studying the Influence of Kindergarten on the Age-Grade Progress of Pupils in New York's Elementary Schools, he found that "The figures indicate that in the elementary schools of New York State, the provision of kindergarten instruction reduced the percentage of failure or nonpromotion in grade one from 20 to six, or by 14 percent.

"More than 80 percent of the kindergarten group made normal or accelerated progress as compared with 58 percent of the first-grade entrants.

"Moreover, the progress data on these 13,730 children showed that retardation or nonpromotion beyond the first grade is much greater for the grade one group than for the kindergarten group."

The ACEI Primary Committee, 1954-56, in "A Study of Feelings of First Grade Teachers Concerning Kindergarten Experiences of Children," (6) found that almost all the teachers replying (98 percent) felt there was a relationship between kindergarten experience and adjustment and progress in first grade. Fast (5) confirmed results of previous studies and found that kindergarteners achieve significantly higher scores on reading tests in first grade than non-kindergarteners.

Readiness Tests. In Table II, the scores on the Metropolitan Readiness Tests are summarized for each group. According to the Metropolitan Tests, 37 percent of the children with high adjustment scores were classified as superior, and two percent as poor risks; only six percent of children having low adjustment scores were classified in the superior group, while 25 percent were classified as poor risks.

In the superior category, the scores on the readiness test correspond closely with ratings of the teachers. Only thirteen, or six percent, of the children in the low adjustment group scored as superior on the Metropolitan Test; and only four, or two percent, of the high adjustment group received scores placing them in category E, classifying them as poor risks. It is interesting to note, however, that 29 percent of the low-adjustment group and 21 percent of the high-adjustment group scored in the average category on the Metropolitan Tests.

The obtained chi-square value reveals that for the subjects used in this study there is a significant relationship at the .01 level between adjustment status and the scores on the readiness test. Hence question two is answered affirmatively.

This finding seems to support Ladd, who, in reviewing correlations between teachers' ratings and teachers' marks, indicated that the correlations which have usually been found are positive and rather high when the traits rated are "desirable," and negative when the traits are "undesirable." (9)

Size of the Family and Ordinal Position in the Family. The percentage of children having high and low adjustment remained nearly the same for each family size and the position in the family, as shown in Tables III and IV.

The highest percentages of both high- and low-adjustment pupils were found to be in the first ordinal position in the family (Table III). The obtained chi-square showed no significant relationship at the .05 level between adjustment status and the ordinal position of the child in the family. Question three is thus answered negatively.

The highest percentage of both high- and low-adjustment pupils was found in the families of three and four children (Table IV). Chi-square calculations for this category indicated that the relationship between family size and adjustment status was not significant at the .01 level — no difference was found in the frequencies of children in high-and low-adjustment categories. Question four is thus answered negatively.

These data tend to support the findings of Beales (3), who found that size of family had little relation to the general personality adjustment of children.

Sheldon (14), however, found that the size of the family and the position in the family, excepting only children, did affect the reading ability of the child. The smaller the family and the earlier the ordinal position in the family, excepting only children, the greater the percentage of good readers. Jones (8) presented evidence from studies of gifted children and gifted men, which have shown a distribution of birth orders strongly favoring the first-born.

The data of Anderson and Kelley (2) suggest that being an only or an eldest child is advantageous to good reading, while having elder siblings is likely to prove disadvantageous. There were fewer only or oldest children in the reading disability group in his study than in the control group, and consequently, more youngest and middle children in this disability group. Findings may be interpreted in various ways, one obvious explanation being that parents are more likely to assist in teaching their only or oldest children to read. This tutoring would be of more value to the brighter than to the duller children.

The question, then, may well be: does adjustment status depend upon the quality of the relationships within the family rather than the size of the family? Several children in the family may provide situations rich in adjustment possibilities, or dominate the situation to such an extent that opportunities are limited.

Socio-Economic Status. The next factor considered was the occupational and educational status of the father (or mother, if she were the breadwinner in the family); as shown in Table V, 28 percent of the children in the high-adjustment group had fathers in the upper-middle class (McGuire-White Scale), while only 12 percent of the low adjustment group had fathers in this category. Only three percent of the children in the high-adjustment group had fathers in the lower-lower class, while 16.3 percent of children in the low-adjustment group had fathers in this category.

The chi-square value obtained indicates that the relationship between adjustment and socio-economic status is significant at the .01 level. Question five is thus answered affirmatively.

While little investigation seems to have been made into the relation between socio-economic status of the home and the school adjustment of the child, some studies relating to success in reading have been made.

Almy (1), in studying children's experiences in relation to success in reading, found no relation between the occupational level of the home and the reading criterion. There was, however, a significant correlation between the occupational level and the responses which the child made in answer to specific questions regarding his experience in the year prior to first grade.

Anderson and Kelley (2) found that reading disability bears little relationship to economic status when a comparison was made between a group of 100 reading disability cases and a control group of 100 other clinic cases. Sheldon (14) found that few good readers come from families where the fathers were employed as semiskilled or unskilled workers, while larger percentages of fathers were in these classifications for average and poor readers. Robinson (13) states that social class affects the range and the type of vocabulary used and understood by pupils, as well as their interests and standards of behavior.

Chronological Age. According to the data presented in Table VI, 51 percent of the high-adjustment group and 49 percent of the low-adjustment group were in the age group between six years and six years-five months. Twenty-four percent of the high-adjustment group and 15 percent of the low-adjustment group were in the age group between six years-six months and six years-eleven months. The most noticeable difference seems to be in the age groups of five years-nine months and five years-eleven months in which 18 percent of the high-achievement group and 31 percent of the low-achievement group were found.

Chi-square calculation indicated significance at the .01 level. For the subjects used, there is a significant relationship between adjustment status and chronological age. This gives an affirmative answer to question six.

Studies made in Florida of the optimum age for entrance into school indicated "a significant difference in reading readiness between the under-age and over-age child, but that the achievement in reading of these two extreme age groups among white children is so slight as to be non-significant by the end of the year." (4) The conclusion was drawn that chronological age alone cannot be used as a criterion for readiness for first-grade learning.

Summary and Conclusions

This study was conducted for the purpose of identifying factors involved in the adjustment of the child in first grade, and attempted to answer the following questions: How do children of high and low adjustment compare in regard to attendance at kindergarten, socio-economic status, ordinal position of the child in the family, size of the family, scores on readiness tests, and the chronological age of the child?

Subjects were selected from the first grades of eight elementary schools in Orange County, Florida. By means of teacher rating, 25 percent of the first-grade children having the highest adjustment scores and 25 percent having the lowest adjustment scores were selected for study. Complete data were secured for 234 children with high-adjustment and 221 children with low-adjustment scores.

Chi-square values were calculated to examine the significance of the relationship between adjustment status and the age of school entrance, attendance at kindergarten, scores on the readiness tests, socio-economic rating, family size and position in the family.

Results of the statistical analysis of the data indicate that for these children there was a significant relationship between high-adjustment status and 1) attendance at kindergarten, 2) scores on reading readiness tests, 3) the socio-economic status of the family, and 4) chronological age. No significant relationship was found between high-adjustment status and the size of the family or the ordinal position in the family. Results were discussed and interpretations of the data offered.

Results from the study indicate the need of further research into the various aspects of the child's adjustment in first grade in order to answer the following questions:

1. What effect does a long or a short period of kindergarten attendance have on first-grade adjustment?
2. Is there relationship between sex and adjustment in first grade?
3. Do the children who make good adjustment in irst grade continue this adjustment through elementary school years?

Such studies involving large numbers of subjects could contribute greatly to further understanding and more effective teaching of first-grade children.

REFERENCES

1. Almy, Corinne. Children's Experiences Prior to First Grade and Success in Beginning Reading (New York: Bureau of Publications, Teachers College, Columbia University, 1949). Quoted in C. W. Hunnicutt and W. J. Iverson, Research in the Three R's (New York: Harper and Brothers, 1958). pp. 48-51.
2. Anderson, M., and Kelley, M. "An Inquiry into Traits Associated with Reading Disability," Smith College Studies in Social Work, II (September 1931), pp. 46-63.
3. Beales, Lester. "A Study of Certain Home Factors and Their Relationship to the Personal Adjustment of Children," School and Society, LII (July 22, 1950), pp. 55-57.
4. Committee Six, Florida Council on Elementary Education. A Study of the Optimum Entrance Age in Florida Schools (1957, processed), p. 9.
5. Fast, Irene. "Kindergarten Training and Grade 1 Reading," Journal of Educational Psychology, XLVII (January 1957), pp. 52-57.
6. Hammond, S. L. "What About Kindergarten?" Childhood Education, XXXIII (March 1957), pp. 314-315.
7. Hildreth, Gertrude, and Griffiths, Nellie. Metropolitan Readiness Tests, Form S (Yonkers-on-Hudson: World Book Company, 1950).
8. Jones, Harold E. "The Environment and Mental Development," Manual of Child Psychology, edited by Leonard Carmichael (New York: John Wiley and Sons, Inc.), pp. 631-696.
9. Ladd, Margaret R. The Relation of Social, Economic and Personal Characteristics to Reading Ability (New York: Bureau of Publications, Teachers College, Columbia University, 1933), p. 6.
10. Lyle, Jean. Questionnaire to Determine Feelings of First Grade Teachers Concerning Kindergarten Experiences of Children, unpublished Master's thesis, University of Tennessee (1955).
11. McGuire, Carson, and White, George D. "The Measurement of Social Status," Human Development (University of Texas: Department of Educational Psychology, 1955).
12. Morrison, J. Cayce. The Influence of Kindergarten on the Age-Grade Progress of Pupils in New York's Elementary Schools (The University of the State of

New York Press, 1938), p. 9. Quoted in Association for Childhood Education, Portfolio on Kindergarten Extension (Washington: Association for Childhood Education, 1941), Section 4.

13. Robinson, Helen M. "Factors Which Affect Success in Reading," Elementary School Journal, LV (January 1955), pp. 263-269.

14. Sheldon, W.D., and Carrillo, L. "Relation of Parents, Home, and Certain Developmental Characteristics to Children's Reading Ability," The Elementary School Journal, LII (January 1952), pp. 262-270.

TABLE I

RELATION OF ATTENDANCE AT KINDERGARTEN TO ADJUSTMENT OF CHILDREN IN FIRST GRADE

Number	Adjustment Status	Type of Kindergarten Experience						Total	
		Summer School		Private		Church Related			
		No.	%	No.	%	No.	%	No.	%
234	High	5	2.1	67	28.6	52	22.2	124	52.99
221	Low	7	3.2	28	12.7	20	9.0	55	24.9

$x^2 = 37.61$

x^2 .01, df = 1 = 6.635

TABLE II

SCORES ON METROPOLITAN READINESS TESTS

Letter Rating on Metropolitan Readiness Test	Adjustment Status			
	High		Low	
	No.	%	No.	%
A - Superior	87	37.2	13	5.9
B - High Normal	75	32.1	29	13.1
C - Average	48	20.5	64	29.0
D - Low Normal	20	8.5	20	27.1
E - Poor Risk	4	1.7	55	24.9
Total	234	100.0	221	100.0

$x^2 = 141.29$

x^2 .01, df=4 = 13.277

TABLE III

RELATION OF ORDINAL POSITION IN FAMILY TO ADJUSTMENT IN FIRST GRADE

Ordinal Position	Number of Children Represented	High Adjustment		Low Adjustment	
		Number	Percent	Number	Percent
1st	157	79	33.8	78	35.3
2nd	146	80	34.2	66	29.9
3rd	86	47	20.1	39	17.6
4th	41	21	9.0	20	9.0
5th	15	5	2.1	10	4.5
6th	5	1	.4	4	1.8
7 or more	5	1	.4	4	1.8

$x^2 = 6.60$ $x^2_{.05,\ df=4} = 9.488$

TABLE IV

RELATION OF FAMILY SIZE TO ADJUSTMENT IN FIRST GRADE

Number in Family	Number of Children Represented	High Adjustment		Low Adjustment	
		Number	Percent	Number	Percent
1 and 2	160	87	37.2	73	33.03
3 and 4	223	114	48.7	109	49.3
5 and 6	57	30	12.8	27	12.2
7 and 8 and 9	15	3	1.3	12	5.4

$x^2 = 6.53$ $x^2_{.05,\ df=3} = 7.815$

TABLE V

RELATION OF PARENTS' SOCIO-ECONOMIC STATUS TO ADJUSTMENT OF CHILDREN IN FIRST GRADE

Adjustment Status of Children	Upper Class		Upper Middle		Lower Middle		Upper Lower		Lower Lower	
	No.	%	No.	%	No.	%	No.	%	No.	%
234-High	12	5.1	66	28.2	81	34.6	68	29.1	7	3.0
221-Low	2	.9	26	11.8	72	32.6	85	38.5	36	16.3

$x^2 = 46.83$ $x^2_{.01,\ df=4} = 13.277$

TABLE VI

NUMBER AND PERCENT OF CHILDREN IDENTIFIED AS HAVING HIGH AND LOW ADJUSTMENT SCORES IN SEVEN AGE GROUPS WITHIN THE FIRST GRADE

Adjustment Status	Below 5-9		6-0 6-5		6-6 6-11		7-0 7-5		7-6 7-11		8-0 8-5	
	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%
High	42	17.9	120	51.3	57	24.4	12	5.1	3	1.3	0	0
Low	68	30.8	109	49.3	34	15.4	6	2.7	1	.5	3	1.4

$x^2 = 13.11$ $x^2_{.01,\ df=3} = 11.345$

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 2, October 1962)

Value Variations in Fourth- and Seventh-Grade Students

PHYLLIS SWARTZ
EDWARD G. NOLAN
KENNETH TILLMAN
University of New Mexico

CONSIDERABLE RESEARCH exists with respect to values and their relationship to such factors as socio-economic position, ethnic origin, sex, amount of education, and religious affiliation of school pupils. In such studies, the question of the stability of a value framework frequently arises. The modification of the value structure with advances in age has intrigued many investigators (1,2,3,4,5,6). For example, how constant are an individual's values over a prolonged period of time? If fluctuations appear, do they obey any regulatory principles? Where fluctuations are apparent, what factors contribute?

The present study is restricted to the changes in value structuring over a definite period of childhood, namely, between the fourth and the seventh grade levels. While it is impossible in a study such as that here reported to isolate the causative factors beyond dispute, certain suggestions will be offered, however, with respect to possible sources of causes, and also an interpretation of differences between the present findings and those of Hawkes (3, 4) which suggested this study.

Hawkes' study was concerned with the changes apparent over the fourth, fifth and sixth grades. To measure such changes, he devised a ninety-item value questionnaire, covering ten basic dimensions of the value continuum. The present study employs this test, utilizing samples of fourth and seventh graders.

The test was administered to several groups in an elementary school and a junior high school in a southwestern city. Two fourth-grade classes were tested in the elementary school and two seventh-grade classes in the junior high school. Precautions were taken to ensure that the schools employed were of comparable socio-economic and religious backgrounds. A total of sixty fourth graders participated. The breakdown consisted of 35 girls and 25 boys. Fifty-four seventh graders were tested, the breakdown being 30 girls and 24 boys.

Procedures

The inventory was completed by all subjects in approximately twenty minutes. The dimensions sampled by this inventory were: Beauty, Comfort and Relaxation, Excitement and Recreation, Friendship, Family Life, Privacy, Physical Freedom, Power and Control, Personal Improvement, and Recognition. Ninety items (nine items per value), representing the ten personal values, were arranged in groups of three. The testees then ranked them one, two, or three, the items being balanced in such a way that no individual dimension was over-represented. The resultant rank order of dimensions represents an approximation only.

Reliability of Hawkes' test was determined through the test-retest method. The median r was .60 with a range from .91 to .42 (3).

Before examining the results of the test, it is necessary to clarify exactly what is meant by the values in question. Following is the interpretation of what each value includes (3):

1. Beauty: Appreciation of beauty and quality in surroundings.
2. Comfort and Relaxation: Primarily physical comfort and relaxation with some reference to mental comfort.
3. Excitement and Recreation: Experiences which give a pleasurable thrill.
4. Friendship: Personal attachments to people outside the home — usually peers.
5. Family Life: Personal attachments to the family unit and the feeling of relation-

ship that exists among its members.

6. Privacy: Being alone and having personal belongings which can be kept from others; freedom from "nosey" adults.
7. Power and Control: Control over others.
8. Physical Freedom: Lack of physical restraint from external sources.
9. Personal Improvement: Improvement of appearance, background, and acquirement of social facility.
10. Recognition: Prestige as such, apart from power and control.

Results

The results of the "Making Choices" test of values were analyzed to determine the differences in values of fourth- and seventh-grade students as reflected by this test. The relative rank of each of the ten values was determined for the fourth-grade girls, the seventh-grade boys, and the composite fourth and seventh grades. The Spearman-Brown correlation technique was employed.

Table I: A correlation of .97 was obtained when comparing the relative ranking of the ten value items of the fourth-grade girls' with the fourth-grade boys' ranking. The boys and girls differed only slightly. The girls ranked family life as the number one value, with friendship number two. The boys reversed these two values with friendship having the highest rating. The only other reversal apparent was with respect to the recognition value and that of power and control. The power and control value was ranked lowest (tenth) by the boys, with recognition ranked ninth. The girls ranked recognition tenth and power and control ninth.

The remaining six items were given the same relative position of importance by both groups.

Table II: The seventh-grade girls correlated .76 with the seventh-grade boys on their ranking of the ten value items. Both groups ranked friendship and family life one and two, respectively. The girls and boys also agreed on excitement and recreation (sixth), privacy (seventh), recognition (ninth), and power and control (tenth).

The boys ranked comfort and relaxation number three, with personal improvement being fifth in importance to them. The girls reversed these two categories.

The largest difference between these groups occurred in the beauty and physical freedom categories. Beauty ranked fourth with the girls, while physical freedom was listed as eighth in importance. The boys of the seventh grade ranked beauty eighth and physical freedom fourth.

Table III: A correlation of .94 was obtained when comparing the composite fourth and seventh grades on the basis of the way they ranked on the ten different categories which compose the value test.

The fourth and seventh grades ranked friendship as the most important value, followed by family life and personal improvement in that order. Both groups ranked beauty fifth, recognition ninth, and power and control tenth.

Excitement and recreation was fourth in importance to the fourth grade, with comfort and relaxation sixth. The seventh grade listed comfort and relaxation fourth with excitement and recreation sixth.

The other difference between these two grades is in their ranking of the privacy and physical freedom values. The fourth grade ranked privacy seventh and physical freedom eighth, whereas the seventh grade reversed the relative position of these two values.

Summary

This study was designed to determine age and sex differences in values as shown by the "Making Choices" Test.

Two fourth-grade classes from an elementary school and two seventh-grade classes from a junior high school were used as subjects. Both schools are located in the same geographical area of a southwestern city. There were 60 fourth-grade students (35 girls and 25 boys) and 54 seventh-grade students (30 girls and 24 boys).

Previous research with this instrument utilizing fourth-, fifth-, and sixth-grade levels is reported by Hawkes (3, 4).

Conclusions

Positive correlations were found for the value rankings of boys and girls in the fourth grade, boys and girls in the seventh grade, and between the total fourth- and seventh-grade groups on this basis.

Although the general agreement between the present study and that of Hawkes (3, 4) is high, there are also some interesting reversals and contrasts which merit consideration. The general agreement is evident from Spearman-Brown correlations for fourth graders (both studies) and for the seventh graders of this study compared to the sixth graders of Hawkes (3, 4). (Fourth graders/Fourth graders = .90; Sixth graders (Hawkes / Seventh graders (present study) = .91.) However, certain other features are worth of consideration.

Sex Differences

Hawkes (3, 4) reported considerable agreement between the sexes with respect to value rankings.

However, it should be pointed out that his groups were composite groups of boys and girls, all three grades included. In the present study, separate correlations were calculated for both sexes in each of the grades under consideration and these are presented in Tables I and II. It should be noted that the agreement with respect to fourth graders is more marked than that with respect to seventh graders (correlations of .97 and .76 respectively). Does this indicate that value discrepancies based on sexual identification become more apparent with time? To what extent is this a function of the approach of puberty? What other factors might be operative in producing such variation over time? When do changes first become apparent?

Critical Ratios

Discrepancies between the present study and that of Hawkes may be due to several factors: Hawkes used sixth graders, the present study seventh graders; also, socio-economic, geographic and cultural factors may be important. The time intervening between these studies and the initial reliability and validity of the scale are other possible sources. There has been little research on value changes of a developmental, longitudinal nature. Such research is essential to a full appraisal of the issues.

REFERENCES

1. Allport, G. W., Vernon, P. E., and Lindzey, G. Study of Values, Manual of Directions, Revised Directions (Boston: Houghton Mifflin Company, 1951).
2. Dahlke, H. O. Values in Culture and Classroom (New York: Harper and Brothers, 1958).
3. Hawkes, G. R. A Study of the Personal Values of Elementary School Children, Ph.D. Dissertation (Ithaca: Cornell University, 1950).
4. Hawkes, G. R. "A Study of the Personal Values of Elementary School Children," Educational and Psychological Measurement, XII (1952), pp. 654-663.
5. Meek, L. H. The Personal-Social Development of Boys and Girls with Implications for Secondary Education (New York: Progressive Education Association, 1950).
6. Witty, P., and Kopel, D. "The Dreams and Wishes of Elementary School Children," Journal of Educational Psychology, XXX (1939), pp. 199-205.

TABLE I

RANK ORDER VALUE PATTERNS FOR FOURTH-GRADE GIRLS AND BOYS

Boys	Girls
1. Family Life	1. Friendship
2. Friendship	2. Family Life
3. Personal Improvement	3. Personal Improvement
4. Excitement and Recreation	4. Excitement and Recreation
5. Beauty	5. Beauty
6. Comfort and Relaxation	6. Comfort and Relaxation
7. Privacy	7. Privacy
8. Physical Freedom	8. Physical Freedom
9. Recognition	9. Power and Control
10. Power and Control	10. Recognition
N = 25	N = 35

= .97

TABLE II

RANK ORDER VALUE PATTERNS FOR SEVENTH-GRADE GIRLS AND BOYS

Boys	Girls
1. Friendship	1. Friendship
2. Family Life	2. Family Life
3. Comfort and Relaxation	3. Personal Improvement
4. Physical Freedom	4. Beauty
5. Personal Improvement	5. Comfort and Relaxation
6. Excitement and Recreation	6. Excitement and Recreation
7. Privacy	7. Privacy
8. Beauty	8. Physical Freedom
9. Recognition	9. Recognition
10. Power and Control	10. Power and Control
N = 24	N = 30

= .76

TABLE III

RANK ORDER VALUE PATTERNS FOR TOTAL FOURTH AND TOTAL SEVENTH GRADES

Total Fourth Grade	Total Seventh Grade
1. Friendship	1. Friendship
2. Family Life	2. Family Life
3. Personal Improvement	3. Personal Improvement
4. Excitement and Recreation	4. Comfort and Relaxation
5. Beauty	5. Beauty
6. Comfort and Relaxation	6. Excitement and Recreation
7. Privacy	7. Physical Freedom
8. Physical Freedom	8. Privacy
9. Recognition	9. Recognition
10. Power and Control	10. Power and Control
N = 60	N = 54

= .94

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 2, October 1962)

The Development of a Parent Attitude Toward Education Scale *

GENE R. MEDINNUS
San Jose State College

WHILE THE importance of the family's attitudes toward education in affecting the child's attitude toward the school and the teacher is well recognized, especially by the classroom teacher, little research has been conducted in this area. Investigations which have obtained some information in this regard have been concerned primarily with social class differences (e.g., 2, 5, 6). In general, research has shown that families of different social levels have differing attitudes toward education and the school. Upper- and upper-middle class parents place great emphasis upon the value of education, while lower-middle and upper-lower class parents regard education as important primarily insofar as it prepares the child for a vocation. Lower-lower class parents often regard the school with suspicion and consider education as essentially unprofitable. Studies of school drop-outs have indicated that the family's attitudes toward education play a major role in the student's decision to leave school.

The data in the research mentioned above have been obtained primarily through interviews with parents, with little attempt to obtain a precise or quantitative estimate of the parents' attitudes. Such information would be useful not only for a broader understanding of family-school relations in general, but specifically, it would aid the teacher in her attempts to understand some of the causes underlying the various attitudes, motivations, and behaviors of her pupils.

The present paper will report data which were collected as part of a larger study dealing with readiness for first grade and first grade adjustment. Since one of the purposes of the study was to identify factors related to a child's success in first grade, some estimate of the parent's attitude toward a number of aspects of schooling and education in general was desired. For this purpose, an Attitude Toward Education scale was developed.

Procedure

The 68 parents, 19 sets of parents of boys and 15 sets of parents of girls, in the present study were all parents of five-year-olds who were subjects in an investigation of school readiness. The location of the field study was a small midwestern community of approximately 5,000 residents. The five-year-old sample itself was not completely representative of the entire population of five-year-olds in the community so far as intelligence and father's occupation are concerned. The mean I. Q. of the five-year-olds was 111.6, according to the Stanford-Binet (Form L) intelligence tests administered by the writer. The socio-economic status of the children's families was approximately evenly divided between the upper-lower and lower-middle classes, according to Warner's Index of Status Characteristics (7). The mean ages of the mothers and fathers were 32.1 and 34.1 years, respectively, and the corresponding mean years of education were 12.0 and 12.2.

Method

The Attitude Toward Education scale included statements covering the following areas: the parent's attitude toward his own educational experiences; the parent's willingness to support the school in matters of discipline, policy, administration and finances; and the parent's evaluation of the importance of education. The forty items comprising the scale are listed:

1. In looking back it seems to me that most of my teachers didn't know what they were talking about.
2. Going to school is a profitable experience. *
3. On the whole schools do a good job in cooperating with parents. *
4. Too much nonsense goes on in classrooms these days.
5. Parents should back up the school in matters of discipline. *
6. Schools should spend more time than they do teaching boys and girls how to be useful citizens.
7. When I look back on my school years, I feel they were well spent. *

Footnotes will be found at the end of the article.

8. My years in school did little to help me in what I'm now doing.
9. The interests I developed while I was in school still provide me with enjoyment. *
10. Many school principals and superintendents are too bossy in their attitude toward children.
11. It seems to me that the public money which is put into schools today is well worth it. *
12. Schools teach a lot of things that don't work out when you actually get on the job.
13. If a child doesn't do well in learning to read it's probably the teacher's fault.
14. Some teachers are so mean you can't blame boys and girls for sassing (or disobeying) them.
15. Some boys and girls are always getting tough breaks in school.
16. On the whole the schools are doing a good job today.*
17. Some teachers ask parents too many questions about how they treat the child at home.
18. Many teachers are more interested in their own social life than they are in educating children.
19. Teachers usually do the best they can in trying to teach what they are supposed to. *
20. There are times when teachers can't be blamed for losing patience with a pupil. *
21. Children should listen to the teacher and do what she says. *
22. It doesn't do any harm for pupils to skip school once in a while.
23. If children listened to everything their teachers told them they'd end up pretty confused.
24. I never could seem to get to school on time.
25. The boys and girls who like school most are usually sissies or teachers' pets.
26. Most schools don't let the parents know enough about what's going on.
27. They're not teaching reading as well as they used to.
28. Many teachers don't explain enough in their teaching.
29. Teachers don't make the pupils behave themselves as well as they should.
30. Many people go into teaching simply because they like to boss others.
31. Schools are the backbone of American Democracy. *
32. Some of my teachers were a source of great inspiration to me. *
33. Most teachers have pets.
34. Most teachers treat the children in their classrooms fairly. *
35. The main reason I can see for going to school is that the law requires it.
36. Most people will learn more by working four years than by going to school four years.
37. The best way to get a good job is to get a good education. *
38. The school is often to blame where students don't like school.
39. Many children would be better off if they left school after the eighth grade.
40. Most teachers teach because they like children.*

The responses to the statements indicating positive attitudes toward education were scored either 3, 2, 1, or 0 depending upon whether the response was "Strongly agree", "Somewhat agree", "Somewhat disagree", or "Strongly disagree", respectively.[1] The scoring of the statements denoting criticism of the school was in the reverse direction, thus the higher scores indicate a more favorable attitude toward education.

The Attitude Education scale was attached to another parent attitude questionnaire which was left with the parents following an evening home visit by the investigator. The questionnaires were returned as they were completed.

In an attempt to obtain another estimate of the parents' attitude toward education and the school, the first grade teacher in whose class twenty-five of the five-year-olds later enrolled was asked to rate the attitudes of these children's parents. For this purpose, a five-point rating scale was used with "3" representing an "average" rating. A rating of "5" denoted an "essentially positive attitude toward the school and education," while a rating of "1" indicated an "essentially negative attitude toward the school and education."

The scores on the test were related to the socio-economic level of the family and to the educational level of the parent.

Results

The split-half reliability of the scale was computed using the Spearman-Brown formula. The correlation coefficient was .90.

With a possible range of 0 to 120, the actual range of scores was from 72 to 118. The means of the mothers' and fathers' scores were 97.74 and 95.34, respectively. This difference is not statistically significant.

When a division of scores was made on the basis of socio-economic level of the subjects, no significant difference was found between the means of the lower-middle versus the upper-lower class parents; these means were 95.84 and 96.70, respectively. Scores on the scale were also analyzed according to the educational level attained by the subjects.

Three categories were used: less than twelve years of education, twelve years of education, and more than twelve years of education. The respective means of the three groups were 92.20, 96.40, and 100.25. The computed F ratio was not statistically significant.

A product-moment correlation of +.26 was obtained between the mothers' versus the fathers' scores.

The mothers' and fathers' scores were correlated separately with the teacher's rating of their attitude toward education and the school. These coefficients were -.07 and +.53, respectively. The latter coefficient was significant at the .01 level.

Discussion

While the differences with regard to the relation between socio-economic level and educational level and the subjects' scores were not statistically significant, several interesting trends were apparent. Scores as related to educational level of subject were in the expected direction. But contrary to discussion in the literature, socio-economic differences were not indicated by the data. The relation between socio-economic level and test score was probably attenuated by the narrow range of socio-economic status of the subjects included in the study. Further study on a wider range would be indicated. It might be, however, that little difference actually exists today between the social classes in the kinds of attitudes which were tapped by the present scale. There is some evidence to suggest that the gap between the social classes is narrowing with respect to child training practices and procedures (1); perhaps this is true also with regard to attitudes.

On the basis of the low correlation between the mothers' and fathers' scores, one can conclude that there is little agreement between parents in their attitudes toward education.[2] Several possibilities for further research are suggested by this finding. First, it might be postulated that this very discrepancy between parents in the value which they attach to education and in the degree to which they support the school's policies produces conflict in the child with regard to his view of the school and the teacher. Second, the relation between the child's attitude and that of the specific parent whose attitudes dominate the home might be examined. A negative attitude by one parent might be relatively unimportant if this parent plays a minor role in influencing the child's attitudes.

The disparity between the coefficients obtained when the parents' scores were correlated with the teacher's rating is difficult to interpret. One would assume that the teacher has greater contact with the mothers and would be better able therefore to assess her attitudes toward the school. The coefficients obtained indicate otherwise. Attempts to verify this finding through further research would seem profitable. Also, an examination of those subjects for whom there is close agreement between the teacher's rating and their Attitude Toward Education scores versus those subjects for whom there is little relationship between the two measures might help identify and define a number of important variables in this area. (For the value of such a procedure, see 3.)

Though the scale needs further development with respect to validation and item discrimination, its present uses for the teacher and the school administration are several. A parent's total score on the scale provides a meaningful estimate of his support for the school and of the importance he attaches to education. Knowledge of the parent's response to various items would be useful to the teacher in identifying specific areas of satisfaction and dissatisfaction. This information might guide the teacher in approaching areas of concern in parent conferences. School administrators might profitably make use of the information also in examining their school policies.

Summary and Conclusions

An Attitude Toward Education scale was developed for the purpose of obtaining an estimate of parents' attitudes toward a number of aspects of schooling and education in general. The scale consists of 40 statements which can be responded to in one of four ways: "Strongly agree", "Somewhat agree", "Somewhat disagree", or "Strongly disagree".

The subjects in the present study, 68 parents — 19 sets of parents of boys and 15 sets of parents of girls, were all parents of five-year-olds who were subjects in an investigation of school readiness. The socio-economic status of the families was approximately evenly divided between the upper-lower and lower-middle classes. The mean ages of the mothers and fathers were 32.1 and 34.1 years, respectively, and the corresponding mean years of education were 12.0 and 12.2.

The scores on the scale were related to the educational level of the parent, the socio-economic level of the family, and to the first grade teacher's rating of the parents' general attitude toward the school and toward the value of education.

The results which emerged from the study may be summarized as follows:

(a) The split-half reliability of the scale was .90.

(b) No statistically significant differences were found when the scores were compared on the basis of education or socio-economic level of the respondents.

(c) A product-moment correlation of +.26 was obtained between the mothers' versus the fathers' scores on the scale.

(d) Coefficients of +.53 and -.07 were obtained between the teacher's rating and the fathers' and mothers' scores, respectively.

Several uses of the scale for the teacher and the school administrator were discussed.

FOOTNOTES

* This research was completed under the sponsorship of the Elizabeth McCormick Memorial Fund of Chicago.

1. These items are marked with an asterisk.

2. This finding supports other data which indicate low inter-parent agreement in various attitudes (4).

REFERENCES

1. Bronfenbrenner, Urie, "Socialization and Social Class Through Time and Space," Readings in Social Psychology, Eleanor E. Maccoby, Theodore M. Newcomb, and Eugene L. Hartley, (Eds.). New York: Holt, 1958. Pp. 400-425.
2. Hollingshead, August B., Elmtown's Youth, New York: Wiley, 1949.
3. Medinnus, Gene R.,"A Note on Interpreting the Pearson r," Journal of General Psychology, in press.
4. Medinnus, Gene R., The Consistency of Inter-parent Agreement on Several Measures, unpublished manuscript.
5. Montague, Joel B., "Social Status and Adjustment in School," Clearing House, 1952, 27, 19-24.
6. Stendler, Celia B., Children of Brasstown, Urbana, University of Illinois Press, 1949.
7. Warner, W. Lloyd, Meeker, M., and Eels, K., Social Class in America, Chicago: Science Research Associates, Inc., 1949.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 2, October 1962)

Behavior Problems of Children as Viewed by Children in the Fifth through Eighth Grades

THOMAS MANGAN and DAVID SHAFER

Rockville Centre, L. I., New York

THE RATING OF BEHAVIOR of children by different professional adult groups has been studied by such authorities as Ellis and Miller (1), Mitchell (3), Stouffer (9), Schrupp and Gjerde (6), Hunter (2), Sparks (7), Stendler (8), and Rivlin (5).

The authors here report a study of children in the fifth through eighth grades to secure their opinions concerning the seriousness of behavior.

With the help of Drs. Richard Clark and Robert Porter (the latter having done a study of student attitudes toward behavior on the secondary and college level (4), the authors composed a list of 47 types of behavior, using the list compiled by E. K. Wickman (10) for a guide. Of the fifty listed by Wickman, three were omitted, viz., heterosexual behavior, masturbation and enuresis.

Procedure

Using the 47 items of behavior on individually dittoed copies (one for each interviewee), the authors contacted 101 children who were in grades five to eight, public, parochial and private schools being represented. Approximately 50 of the interviewees were female and the same number were male.

Upon approaching the child, the interviewer would ask the child's cooperation in giving his views on the seriousness of different types of behavior. If the child were agreeable to giving his time for the survey, he was handed a two-page form, listing the 47 types of behavior with four columns for choice. They were: very serious, serious, not serious and trivial. The interviewer then stayed with the child, explaining the categories of choice and also the types of behavior if there was the slightest doubt in the child's mind concerning them.

The results were correlated with the traditional and modern teachers' lists in Wickman's study. The authors also correlated the study with the original Wickman list made by the mental hygienists.

Dr. Robert Porter's recent study, which was mentioned earlier (4), was also used. The authors correlated their study with the results Porter obtained from high school and college seniors.

Results

Children tend to acquire the attitudes of their teachers toward certain behavioral types. There was a +.746 correlation with the traditional teachers. It may further be observed that students retain their attitudes on the seriousness of certain types of behavior through high school and even through college.

REFERENCES

1. Ellis, D. B., and Miller, L. W. "Teachers' Attitudes and Behavior Problems," Journal of Educational Psychology (October, 1936), pp. 501-511.
2. Hunter, E. C. "Changes in Teachers' Attitudes Toward Children's Behavior over the Last Thirty Years," Mental Hygiene, III (January, 1957), pp. 3-11.
3. Mitchell, John C. "A Study of Teachers' and Mental Hygienists' Ratings of Certain Behavior Problems of Children," Journal of Educational Research, XXXVII (December, 1942), pp. 292-307.
4. Porter, Robert M. "Student Attitudes Toward Child Behavior Problems," Journal of Educational Research (May, 1959), pp. 349-352.

RESULTS OF INTERVIEWS

Ranking by the Various Groups	R_1	R_2	R_3	R_4	R_5	R_6
1. Stealing	1	1	1	8	0	1
2. Skipping school	2	4	3	17	6	17
3. Cheating	3	7	15	24	1	3
4. Destroying school material	4	8	6	21	0	2
5. Temper tantrum	5	11	11	15	9	12
6. Lying	6	3	8	19	3	4
7. Obscene materials	7	2	13	32	4	7
8. Swearing	8.5	13	37	45	2.5	15
9. Smoking	8.5	16	34	46	24.5	32
10. Disobedience	10	9	9	38	10	9
11. Defiance	11	5	5	29	6	5
12. Lack of interest	12	12	19	23	2	19
13. Impoliteness	13	14	20	40	4	13
14. Thoughtlessness	14.5	35	42	36	29.5	42
15. Bullying	14.5	6	2	5	9.5	6
16. Imaginative lying	16.5	39	46	37	10.5	35
17. Resentful	16.5	26	10	7	9.5	20
18. Unreliable	18	10	7	20	11	10
19. Slovenly in appearance	19	31	39	30	9	37
20. Easily discouraged	20	20	21	14	1	11
21. Selfishness	21	21	16	25	8	18
22. Inattention	22.5	23	25	28	14.5	36
23. Quarrelsomeness	22.5	24	17	26	1.5	14
24. Domineering	24	30	18	16	8	22
25. Inquisitiveness	25.5	41	44	41	13.5	44
26. Silliness	25.5	36	27	35	3.5	40
27. Suspiciousness	27	34	32	4	14	28
28. Careless in work	28	22	31	39	14	33
29. Overcritical	29	42	29	13	8	21
30. Stubbornness	30	29	35	34	10	39
31. Disorderliness	32.5	18	26	42	9.5	23
32. Interrupting	32.5	40	45	44	1.5	45
33. Sullenness	32.5	32	27	22	2.5	29
34. Laziness	32.5	15	29	33	17.5	26
35. Suggestible	35	25	22	12	13	27
36. Tardiness	36	27	41	43	6	43
37. Tattling	37	43	43	31	12	30
38. Physical coward	38	28	28	18	0	31
39. Dreaminess	39	38	36	10	4	38
40. Unsocial	40.5	37	12	1	8.5	16
41. Unhappy	40.5	19	4	2	22.5	8
42. Whispering, note writing	42	44	47	47	4	46
43. Sensitiveness	43	45	30	9	1	41
44. Fearfulness	44	33	23	3	13	24
45. Nervousness	45	17	14	11	9	25
46. Restlessness	46	46	40	27	1	47
47. Shyness	47	47	32	6	0	34

Legend:

R_1 — The author's
R_2 — Wickman's Traditional Teachers
R_3 — Wickman's Modern Teachers
R_4 — Wickman's Mental Hygienists
R_5 — Porter's High School Seniors
R_6 — Porter's College Seniors

Using Spearman's formula, we obtained the following correlations:

With the Wickman traditional teachers, +.746
With the Wickman modern teachers, +.457
With the Wickman mental hygienists, −.196
With the Porter high school seniors, +.712
With the Porter college seniors, +.610

5. Rivlin, Harry N. *Educating for Adjustment* (New York: Appleton-Century-Crofts, 1936), p. 254.
6. Schrupp, Manfred H., and Gjerde, Clayton M. "Teacher Growth in Attitudes Toward Behavior Problems of Children," *Journal of Educational Psychology*, XLIV (April, 1953), pp. 203-214.
7. Sparks, J. N. "Teachers' Attitudes Toward the Behavior Problems of Children," *Journal of Educational Psychology*, XLIII (May, 1952), pp. 284-291.
8. Stendler, Celia B. "How Well do Elementary School Teachers Understand Child Behavior?" *Journal of Educational Psychology*, XL (December, 1949), pp. 489-498.
9. Stouffer, George A. W., Jr. "Behavior Problems of Children as Viewed by Teachers and Mental Hygienists," *Mental Hygiene*, XXXVI (April, 1952), pp. 271-285.
10. Wickman, E. K. *Children's Behavior and Teachers' Attitudes* (New York: The Commonwealth Fund, Division of Publications, 1928).

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 2, October 1962)

The Laboratory School: An Historical Perspective

POSE M. LAMB
Ball State Teachers College

LABORATORY EXPERIENCES for students and demonstration teaching by "expert" teachers have long been considered valuable components of most teacher-education programs. There is, however, some disagreement as to exactly when and where laboratory schools were first established in order to provide facilities for the performance of these functions.

Williams credits the Franciscan Friars in what is now the state of New Mexico with providing student teaching experiences for prospective teachers as early as 1600.* (9, p. 2) Cubberly indicates that the Abbe de la Salle was the first to establish normal schools, and that the origins of such institutions are French. He states that the first normal school anywhere was established at Rheims, France, in 1685. (2)

Duke Ernest of Gotha willed a sum of money to establish schools in order that prospective teachers might, "through practice, learn that for which they will, in the future, be employed." (6) Basedow, a member of the religious order called the Pietists, founded a school at Dessau, Germany, in 1774, where teachers were trained, and as part of that training, worked directly with children. (9, p. 1)

Pestalozzi was instrumental in the establishment of two schools where teachers were given the opportunity to observe and instruct children. The first such school was established at Burgdorf, Switzerland, and was superseded by the more famous school at Yverdon, which Pestalozzi operated from 1805 to 1825. (9)

Hall writes that Herbart founded a practice school at the University of Konigsberg soon after joining the faculty of that institution in 1809. Karl Volkmar Stoy, a pupil of Herbart's, established his own seminary and practice school at Jena, in Germany, in 1843. Stoy was succeeded by Professor W. Rein, who continued to operate the school according to Herbartian principles. (4)

A seminary and practice school were established in Leipzig, Germany, in 1857. Professor Zuller, also a follower of Herbart's, operated this school until his death in 1883. (4)

Herbart and Pestalozzi, among others, directly influenced American concepts of teacher education. Good writes: "It will not be denied that the normal school was an academy with reviews of the common branches and a few professional studies and skills added. It would have been purely American if it had not incorporated Pestalozzi's theory and practice into its professional work." (3)

Herbart's influence on teacher education in the United States, and on laboratory schools in particular, came through the work of F. W. Parker, and the organization now known as the National Society for the Study of Education. (3, p. 13)

European educators, for the most part, conceived the laboratory school as fulfilling one or both of the following functions: these schools furnished places where prospective teachers might practice with the specific guidance and supervision of an experienced teacher, and/or they provided "stages" on which currently approved methods and techniques of teaching could be illustrated. The laboratory school came to the United States with these functions, student teaching and demonstration, well established.

Laboratory schools have been important adjuncts of most teacher education institutions. The titles assigned to the schools give some indication of their primary function. Perodin writes: "From its earliest beginnings, a distinctive feature of teacher education has been the use of an actual school for children. Names given to these children's schools have varied according to their purposes. In the early days of teacher education in this country, they were named "model schools." Later, some were designated as "practice schools," then "training schools," and more recently, the term “laboratory schools” has come into usage.” (6, p. 1)

---

* References will be found at the end of the article.

The first model or practice schools in the United States were privately owned and operated, as were the teacher training institutions with which they were associated. An example was “Mother Seaton's Teacher Training School,” located in Emmitsburg, Maryland, which was founded in 1808. (6, p. 2)

Samuel Hall opened his school at Concord, Vermont in 1823. (8) The school was closed in 1830. (8) The first publicly supported normal school was opened at Lexington, Massachusetts on July 3, 1839, with Cyrus Pierce as principal. (6, p. 2) Good reports that in 1860 there were twelve state normal schools, and that following the Civil War, approximately twenty-five new teacher training institutions were established each year. (3, pp. 213-214)

Barnard's findings indicate that model or practice schools, and occasionally both, were established along with, or soon after, the opening of most of these schools. (1) David Page opened a normal school at Albany, New York, in 1845. One room of the school was designed to serve the “model” or demonstration function, and the other was for practice teaching. A supervising teacher was in charge of each room. (6, p. 4)

Edward Sheldon required prospective teachers to spend one year in observation and practice teaching in his “Oswego, New York Primary Training School.” This was a much longer period of time than was typically devoted to such activities. (6, p. 4) A United States Commissioner of Education Report indicates that by 1875, “forty-seven of the sixty-seven state normal schools had laboratory schools attached to them.” (6, p. 6)

The laboratory school became an active participant in the so-called scientific movement in education which, Judd states, began in the decade between 1890 and 1900. (5) In fact, a few of the outstanding “research centered” schools were established prior to the beginning of the scientific movement. F. W. Parker became the principal of the Cook County Normal School in 1883. Experimentation and “investigation on the work of teaching” were very important functions of this laboratory school. (7, p. 88)

An experimental school was opened concurrently with Teachers College in New York City in 1887. (6, p. 6) John Dewey founded the laboratory school at the University of Chicago in 1896. Rugg writes: “Its aim was to further the application of scientific concepts and methods to the conduct of school work.” (7, p. 94)

Teachers College established a second laboratory school in 1899. Speyer School was the setting for numerous experimental studies and research projects. (6, p. 9) Another experimental school was established by J. L. Meriam at the University of Missouri. (6)

In 1917, Lincoln School was established as the third laboratory school associated with Teachers College. There was no practice teaching at Lincoln School and much emphasis was placed on “. . . trying to discover new and better materials of instruction and improved methods of organization and teaching.” (7, p. 10) Ohio State's University School was opened in 1932, with Dr. Laura Zirbes as director of the elementary school and Dr. Rudolph Lindquist as principal. Experimentation continues to be an important function of this school.

Lincoln School and Horace Mann School were merged in 1941. In 1948, the laboratory schools were abandoned, and the Horace Mann-Lincoln Institute of School Experimentation was organized. (10)

Summary

While there is some controversy over the location and date of establishment of the first laboratory school, it is clear that persons concerned with the education of teachers have, for several centuries, recognized the value of providing facilities for guided, supervised student teaching, and the importance of the demonstration of approved educational methods.

Early laboratory schools typically served one or both of the functions suggested above, and their titles, “model schools,” “practice schools,” “demonstration schools,” give some indication of the purpose or purposes for which the schools were established. As the schools assumed more functions, more general titles, such as “campus schools” or “laboratory schools” were more frequently used. It would be a mistake to assume that the advent of relatively free, publicly supported normal schools in the United States meant that most beginning teachers had the benefits of some professional education, including supervised work with children. However, particularly during the period immediately following the Civil War, there was a great increase in the number of normal schools established, and more prospective teachers were given something more than a review of the “common branches” as part of the preparation. Laboratory schools were integral, even central parts of these normal schools.

Although the so-called “scientific movement” in education met with initial resistance in normal schools, a few laboratory schools were soon outstanding participants in educational research. The leadership in educational research provided by Dewey and Judd at Chicago and J. L. Meriam at the University of Missouri is cited as being of great significance by most writers concerned with the growth and development of the laboratory school.

REFERENCES

1. Barnard, Henry. Normal Schools and Other Institutions, Agencies, and Means Designed for the Professional Education of Teachers (Hartford: Case Tiffany and Company, 1851), pp. 69, 247, 251, 261.
2. Cubberly, E. P. The History of Education (Boston: The Houghton Mifflin Company, 1920) p. 744.
3. Good, H.G. A History of American Education (New York: The MacMillan Company, 1956), p. 12.
4. Hall, J.W. "Professor Rein's Practice School at Jena, Germany," National Education Association, Journal of Addresses and Proceedings, 1899 (Chicago: University of Chicago Press, 1899), pp. 882-884.
5. Judd, Charles H. "The Influence of Scientific Studies in Education on Teacher Training Institutions," Peabody Journal of Education, II (May 1925), p. 296.
6. Perrodin, A. F. "The Development of Laboratory Schools," Functions of Laboratory Schools in Teacher Education (Lock Haven, Pennsylvania: The Association for Student Teaching, 1955), p. 1.
7. Rugg, Harold. "Curriculum Making in Laboratory Schools, " National Society for the Study of Education, 26th Yearbook, Part I, Curriculum Making, Past and Present (Bloomington, Illinois: Public School Publishing Company, 1926), p. 88.
8. Stone, M.S. "The First Normal School in America," Teachers College Record,XXIV (May, 1923), p. 270.
9. Williams, E.I.F. The Actual and Potential Use of Laboratory Schools (New York: Bureau of Publications, Teachers College, Columbia University, 1942), p. 2.
10. Young, Bernard Jae. An Exploratory Study of Some Contributions Campus Elementary Schools May Make to Processes of Curriclum Development in Public Schools (unpublished doctoral dissertation, University of Illinois, 1957), p. 126.

JeR Field News

A SYSTEM MODEL FOR MEASURING ORGANIZATIONAL EFFECTIVENESS is recommended by Amitai Etzioni in a recently published article entitled: "Two Approaches to Organizational Analysis: A Critique and Suggestion." The author points out that a basic methodological error often occurs if private or public goals of cultural entities are compared with social systems—the mistake being that objects which are not on the same level of analysis are being compared; that is, the ideal with the real. The model proposed by the author theoretically is more powerful and avoids certain value judgments; however, it tends to be more demanding and expensive for the researcher. A reprint of this article may be obtained from Bobbs-Merrill Company, College Division, 1720 East 38th Street, Indianapolis 6, Indiana.

OPPORTUNITY FOR GRADUATE STUDY OR RESEARCH in forty-six countries for the academic year 1936-64 will be made to more than 800 American students under the Fulbright-Hays Act of 1961. Three types of grants are available:

1. Full U. S. government grants that provide round-trip transportation, maintenance, and tuition and books.
2. Joint U. S.-other government grants that provide a travel award from the U. S. government in conjunction with foreign government grants for tuition and full or partial maintenance.
3. Travel-only grants that supplement scholarships received from foreign governments, universities, or private donors.

Further information and applications will be accepted until November 1, 1962. Information may be obtained from the Information and Counseling Division, Institute of International Education, 800 Second Avenue, New York 17.

EDUCATIONAL VIEWPOINTS AND POPULAR PHILOSOPHIES OF EDUCATION definitely are regional according to a report published in the University of Alberta *Research Newsletter*. Four thousand opinions from four regions in the United States and one region in Canada were gathered. The major findings are summarized as follows:

The American Midwest assumed what might be termed a middle-of-the-road position; this position seemed to be a pivotal point around and across which major deviations were found.

The West deviated in the direction of more socialization in schools; The South emphasized personal development, particularly physical; the East favored greater attention to moral and vocational training; and Canada, the greatest deviant, placed significantly greater emphasis upon intellectualism, aesthetic appreciation, and world citizenship.

THE EDUCATIONAL AND SOCIAL HANDICAPS RESULTING FROM CHILDHOOD HEART DISEASE OR RHEUMATIC FEVER are more likely to keep young people from adjusting normally to jobs and other responsibilities of adult life than the purely physical effects of their illness. This is the result of a seven-year pilot project in vocational counseling recently published by the American Heart Association. Two hundred and twnety-eight youngsters of the study group were followed from the 8th- or 9th-grade levels to one year after graduation from high school. They were children with actual or potential heart handicaps, but of a nature judged not so severe as to rule out an employable future. Heart disease and rheumatic fever imposed interrruptions in the usual flow of childhood experiences, thus seriously interferring with school progress and personality development. It was found that children in these medical categories need more help than normal children in adjusting to school requirements, including remedial work as well as vocational guidance in seeking out jobs.

A RESEARCH PROJECT IN EXECUTIVE DECISION MAKING was conducted jointly by the Harvard Business School and Educational Testing Service during the past summer. Business executives were offered an opportunity to participate in an unusually complex and realistic business game developed by the staff of the ETS Executive Study and members of the Harvard Business School faculty. Each participant was assigned to a role as a vice-president for manufacturing, vice-president for sales, vice-president for research and engineering, or controller of a business modeled closely on an existing company. He was asked to deal with a wide sample of operating problems and special decisions in such areas as research and development, expansion of product and market, purchase of other companies, and interpretation of predictive data. Observations of their behavior will be analyzed to determine variables that are significant for effective decision making.

A PUBLIC SCHOOL TEACHER IS LIKELY TO REPRESENT AN INVESTMENT OF 300,000 DOLLARS according to Professor Bernard H. McKenna of the Department of Educational Administration of Teachers College, Columbia University. Professor McKenna recommends that schools that want to hire a really good teacher should be prepared to spend another $3000 toward recruitment. The $300,000 is a long-term investment of professional service of twenty to thirty years. One percent of that amount, or $3000, represents what it should be worth to a school to insure top quality personnel.

* * * * *

Readers are invited to send items of research interest to John Schmid, Professor, College of Education, University of Arkansas, Fayetteville, Arkansas.

JACOB O. BACH, Head, Educational Research Bureau, Southern Illinois University, Carbondale, Illinois. Critique Editor.
WILLIAM A. BROWNELL, Dean, School of Education, University of California, Berkeley, California. Education Psychology, Learning, and Teaching.
ROBERT A. DAVIS, Professor of Educational Research, George Peabody College for Teachers, Nashville, Tennessee. Evaluative Techniques.
GUY D. McGRATH, Dean, College of Education, Arizona State University, Tempe, Arizona. Secondary Education.
HAROLD E. MITZEL, Director, Office of Research and Evaluation, Division of Teacher Education, College of the City of New York, New York City. Teacher Personnel.
WILLARD C. OLSON, Dean of College Education, University of Michigan, Ann Arbor, Michigan. Child Development.
JOHN SCHMID, Professor, College of Education, University of Arkansas, Fayetteville, Arkansas. Field News Editor.
WILSON B. THIEDE, Professor of Education, University of Wisconsin, Madison, Wisconsin. Book Review Editor.
ARTHUR E. TRAXLER, Executive Director, Educational Records Bureau, New York 32, New York. Educational Measurements.

Associate Editors

HAROLD M. ANDERSON, Associate Professor of Education, University of Colorado, Boulder, Colorado.
VERNON E. ANDERSON, Dean, College of Education, University of Maryland, College Park, Maryland.
RUDYARD K. BENT, Professor of Education, University of Arkansas, Fayetteville, Arkansas.
EMMETT ALBERT BETTS, Research Professor, University of Miami, Coral Gables, Florida
LAWRENCE P. BLUM, Assistant Director, Student Personnel Services, University of Wisconsin–Milwaukee, Milwaukee 3, Wisconsin.
J. MURRAY LEE, Chairman, Department of Elementary Education, Southern Illinois University, Carbondale, Ill.
D. WELTY LEFEVER, Professor of Education, University of Southern California, Los Angeles, California.
L. JOSEPH LINS, Professor and Coordinator, Institutional Studies, University of Wisconsin, Madison, Wisconsin.
S. E. TORSTEN LUND, Associate Professor of Education, University of California, Berkeley, California.
WILLIAM A. McCALL, Lecturer in Education, University of Miami, Coral Gables, Florida.
DELYTE W. MORRIS, President, Southern Illinois University, Carbondale, Illinois.
VALWORTH PLUMB, Chairman, Division of Education and Psychology, University of Minnesota (Duluth Branch), Duluth, Minnesota.
DAVID G. RYANS, Rand Corporation, Santa Monica, California.
LESTER B. SANDS, Professor of Education, University of California, Santa Barbara, Goleta, California.
MANDEL SHERMAN, President, and Consultant, Emotional Health Association, Los Angeles, California.
MILDRED B. STANTON, Consultant, Special Education, State Department of Education, Hartford 15, Conn.
W. W. THEISEN, Professor of Education, Marquette University, Milwaukee, Wisconsin.
M. R. TRABUE, Professor of Higher Education, University of Kentucky, Lexington, Kentucky.
J. WAYNE WRIGHTSTONE, Director, Bureau of Educational Research, Board of Education of City of New York, Brooklyn, New York.

Editors Emeriti

E. J. ASHBAUGH, Dean Emeritus, School of Education, Miami University, Oxford MR50, Ohio.
HARRY J. BAKER, 19050 Wiltshire, Lathrup Village, Michigan.
FOWLER D. BROOKS, Director Emeritus, Department of Psychology, DePauw University, Greencastle, Ind.
LEO J. BRUECKNER, Professor Emeritus, University of Minnesota, Minneapolis, Minnesota.
B. R. BUCKINGHAM, 33 Chatham Circle, Wellesley Hills, Massachusetts.
EDGAR A. DOLL, Director Emeritus, Devereux Schools, Devon, Pennsylvania.
HARL R. DOUGLASS, Director Emeritus, School of Education, University of Colorado, Boulder, Colorado.
E. DUNCAN GRIZZELL, Dean Emeritus, University of Pennsylvania, 640 Maxwelton Court, Lexington, Ky.
ERNEST HORN, Professor Emeritus of Education, State University of Iowa, W307 East Hall, Iowa City, Iowa.
EARL HUDELSON, Professor Emeritus of Education, West Virginia University, Morgantown, West Virginia.
A. M. JORDAN, Professor Emeritus, University of North Carolina, Chapel Hill, North Carolina.
ALFRED S. LEWERENZ, Director Emeritus of Evaluation and Research, 4107 Holly Knoll Drive, Los Angeles, California.
A. R. MEAD, Director Emeritus of Educational Research, University of Florida, 1719 N. W. 6th Ave., Gainesville, Florida.
M. J. NELSON, Dean Emeritus, Iowa State Teachers College, 6033 West Orange Drive, Glendale, Arizona.
JESSE B. SEARS, Professor Emeritus, Stanford University, Stanford, California.
CARLETON W. WASHBURNE, Chairman Emeritus, Department of Education, Brooklyn College, Brooklyn, New York.

Staff

JEAN WILLIAMS, Supervisor of Editorial Services. **CLARENCE A. SCHOENFELD**, Editorial Associate. **ARNOLD CAUCUTT**, Business Manager.

The Journal of Educational Research

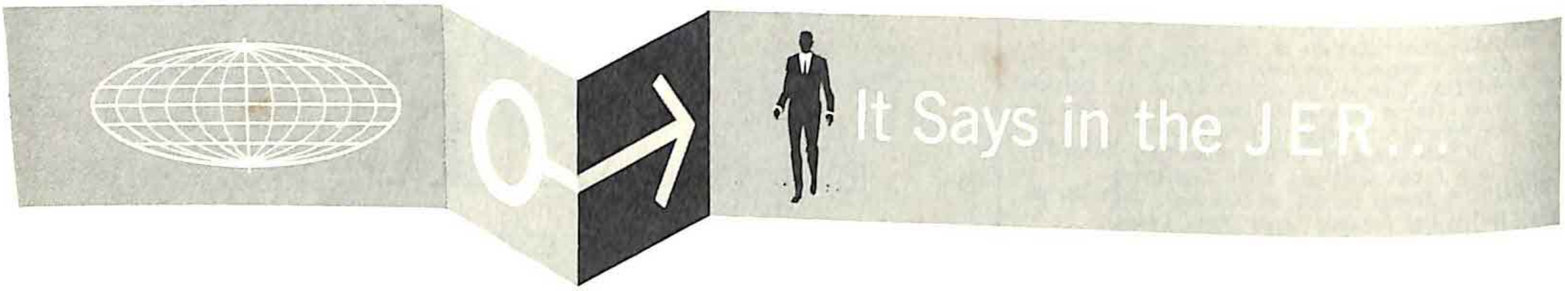
Dedicated to the Scientific Study of Education

VOLUME 56 **NOVEMBER 1962** **NUMBER 3**

It Says in JER ---------- 114
Among Those Present ---------- 116
How Much Are Students in the Public Schools Getting from Their School Work?—Doyne M. Smith ---------- 117
Development of and Reasons for Vocational Choices of Adolescents Through the High-School Years—Marvin Powell and Viola Bloom ---------- 126
The Relation Between Junior High-School Students' Problems and Parental Attitude Toward Child Rearing and Family Life—Michael Zunich ---------- 134
An Experiment in Teaching Elementary School Methods via Closed Circuit Television—Herbert P. Rumford ---------- 139
Bilingualism in Southwest Louisiana—Harold D. Love ---------- 144
A Study of the Relationship Between Parental Acceptance and the Academic Achievement of Adolescents—Janice M. Barwick and Dugald S. Arbuckle ---------- 148
Teachers' Ratings of High-School Students on "Likability" and Their Relation to Measures of Ability and Achievement—J. R. Williams and Walter W. Knecht ---------- 152
Grade Placement of High-School Biology—H. Seymour Fowler 156
Personality Ratings of High-School Students—Donald A. Leton ---------- 160
An Investigation of the Relationship of Some Pre-Adolescent Developmental Factors to Adolescent Social Adjustment—Israel Woronoff ---------- 164
Field News—John Schmid ---------- 167

The Journal of Educational Research is published at Madison, Wisconsin, monthly 10 times a year (the May–June and July–August issues being combined) by Dembar Publications, Box 1605, Madison 3, Wisconsin. The subscription price is $7.50 a year plus $1 a year postage for all subscriptions outside the continental United States. Single copies, $1. Address all communications to Dembar. Second class postage paid at Madison, Wisconsin.

COPYRIGHT 1962 BY DEMBAR PUBLICATIONS, MADISON 3, WISCONSIN



It Says in the JER...

SOME DISTURBING CONCLUSIONS are arrived at by Powell and Bloom in a study of the development of and reasons for vocational choices in high-school students: 1) youths are frustrated in the intelligent selection of a vocation because of lack of knowledge of vocational fields, 2) adolescents are not really aware of their motives for selecting vocations, 3) they have had little effective vocational guidance, and 4) student goals are immature and peurile.

STUDENTS' PROBLEMS are related to parental attitudes toward child rearing, Zunich reports from a study of junior-high pupils in Texas.

IN A SEARCH FOR BETTER MEANS of predicting school achievement, Williams and Knecht get a high correlation between teachers' ratings of high-school students on "likability" and student academic ability scores. The question remains whether such ratings represent very little more than imperfect reflections of academic performance or whether trait ratings do indeed tap personality variables that affect academic achievement. The researchers offer a predictive equation for further testing.

"ALL THINGS CONSIDERED, how much are pupils in your school getting out of school work?" Smith put this question to teachers, parents, and students in 20 of Georgia's 198 school systems. With the exception of sixth- and eighth-grade students, a higher percentage of each respondent group believed that students are getting either somewhat less or a great deal less out of school work than there were those who indicated that they were getting about all they could. Yet there is nothing in Smith's data to indicate why this should be so.

WHEN YOU STUDY the social adjustment of pre-adolescents, you must inject physical-development factors into the problem, Woronoff says in reporting on exploratory studies at the University of Michigan.

BARWICK AND ARBUCKLE report a significant relationship between parental acceptance and the academic achievement of adolescents.

WE CAN PLACE BIOLOGY in either the ninth or tenth grade and find that our students will achieve equally well in either grade, Fowler says.

WHEN TEACHERS ARE ASKED TO judge student personalities, do you get a bona fide personality rating or do you get an academic-ability halo effect? Leton sheds some light on the problem by correlating students' scores on CPI scales with teachers' rating of the same students. He finds significant relationships for 17 of the 18 scales. When considered in reference to other findings, his research suggests that teacher ratings are most useful in rating personality traits associated with scholastic achievement and sociability.

IN A TEST OF TEACHING ELEMENTARY SCHOOL METHODS via closed circuit television, Rumford found televised instruction and TV observation of classroom procedures as effective as conventional instruction and direct observation.

Among Those Present

DOYNE M. SMITH holds BS and MS degrees from Arkansas and an EdD from Colorado. He has held school administration posts in Arkansas and Missouri. At the University of Georgia he is chairman of the department of educational administration and director of the bureau of educational studies and field services.

HERBERT P. RUMFORD is associate professor of education and psychology at Kansas State College of Pittsburg. He holds BA and ThB degrees from Eastern Baptist Theological Seminary, an MA from Minnesota, and an EdD from Colorado.

DONALD A. LETON, a Minnesota PhD, is assistant professor of education at UCLA. He was formerly on the staff of the Minnesota Department of Education.

MICHAEL ZUNICH holds degrees from Ohio State, Akron, and Florida State. He is an associate professor of research in child development at Texas Woman's University.

ISRAEL WORONOFF holds three degrees from the University of Michigan, has taught at Flint Junior College, State College of Minnesota, and the University of Michigan, and is now professor of education at Eastern Michigan University.

J. ROBERT WILLIAMS, an Illinois PhD, is school psychologist at Kankakee, where W. W. KNECHT, an Illinois MA, is superintendent of schools.

HAROLD D. LOVE is at present on sabbatical leave from F. T. Nicholls State College, Thibodaux, Louisiana, while working on his doctorate at Colorado State. He holds both BS and MS from Mississippi Southern College.

JANICE BARWICK is a counselor at Plainview High School in New York and an instructor in guidance at Hofstra. She holds degrees from Pembroke and Boston. DUGALD S. ARBUCKLE, a Chicago PhD, is director of counselor education and professor of education at Boston. A former RCAF officer, he is the author of five books in the field of counseling.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 3, November 1962)

How Much Are Students in the Public Schools Getting From Their School Work?

DOYNE M. SMITH
University of Georgia

CRITICS of public education appear to be getting more space in popular magazines and newspapers than professional educators and researchers. Criticisms boil down to the contention that public-school students are not getting as much out of school work as they should. If there is any basis for this contention, a poll of teacher, parent, and student opinions should reveal some of the reasons.

Purposes of the Study

The purposes of this study were twofold: 1) to poll teachers, parents, and students to get their opinions on how much students are getting out of their school work, and 2) to analyze the student, teacher, and parent opinions about selected factors and/or conditions which should support their expressed beliefs.

The Illinois Opinion Inventories by Hand, Finlay and Dolio, published by the World Book Company, were selected as the polling instruments. These consist of five inventories, each separately designed for upper elementary-school pupils, secondary-school pupils, teachers, patrons, and non-patrons. With the exception of the non-patron series, all of these were used.

The Sample

Twenty of Georgia's 198 school systems participated in the study. They were geographically representative of the state and, excluding metropolitan areas, were reasonably representative of the state in size, number of schools, curricula, and socio-economic factors that affect schools.

The sample included 20 high schools, 10 schools with grades one through twelve, and 78 elementary schools. All of the teachers, all of the students of the sixth, eighth, and twelfth grades, and the parents of these students participated in the study. Parents having children in more than one of these grades were asked to complete the opinionnaire about the school or grade in which their oldest child was enrolled. Participants totaled 8894. Of these, 1126 were teachers (691 teachers of the sixth and eighth grades, and 427 teachers of the twelfth grade), 2999 parents (1703 parents of sixth- and eighth-grade students, and 1296 parents of twelfth-grade students), and 4779 students (1927 sixth-grade, 1616 eighth-grade, and 1236 twelfth-grade students).

Administering the Opinionnaire

Preparation for the survey began in September. The purpose of the study was discussed in faculty meetings, student groups, Parent-Teacher Association meetings, civic clubs, and farm bureau meetings. The University of Georgia's News Bureau prepared periodic news releases for local radio stations serving the area of each school system involved.

After a three-months orientation period, the Opinion Inventories were given to participants. All sixth-, eighth-, and twelfth-grade students were brought together at the same time in centrally located schools, and they completed the inventories under the superivison of Parent-Teacher Association members. The completed forms were forwarded directly to the Director, Bureau of Educational Studies and Field Services, University of Georgia, College of Education, who coordinated the survey. The same procedure was followed with teachers. The Inventories of Parent Opinion were mailed directly to the parents, and they, in turn, forwarded the completed forms to the Director of the Bureau. All respondents were assured of complete anonymity.

The basic question of the portion of the study reported here is: All things considered, how much (are pupils in your school, is your child, are you) getting out of school work? Possible responses were: "no opinion," "about all (they, he, I) could get," "somewhat less than (they, he, I) could get," and "a great deal less than (they, he, I) could get."

Each opinionnaire contains approximately seventy-five questions. However, only selected questions were used in this inquiry. One, as indicated above, was selected as the basic question. The others were selected on the assumption that an analysis of the responses to them should, if respondents were consistent in indicating their beliefs, reveal some of their reasons for responding as they did to the basic question. These questions, with few exceptions, are stated so that each group of respondents answers the same question. Each question has possible responses arranged in a continuum similar to that of the basic question cited above, which is designed to show the respondent's degree of satisfaction or dissatisfaction.

For comparative purposes, responses to the basic question and selected related questions are reported below. Seven of these are directly related to the schools' academic programs. Questions pertaining to student activities and discipline were included because of the oft-heard complaints that students participate in too many "extra-curricular" activities and that the schools "don't teach discipline any more." In addition to these questions, a distribution of students' grades has been included. Responses to a question having to do with general satisfaction with schools are also reported in order to compare the respondents' overall satisfaction with the degree of satisfaction shown for specific factors and/or conditions.

Analysis

Each question used, beginning with the basic question, is listed below, and immediately following it is a summary of the responses to that question. For analysis, sixth- and eighth-grade teacher and sixth- and eighth-grade parent respondent groups are combined.

Responses are also combined when possible in order to compare the percent of respondents who are satisfied with those who indicated any degree of dissatisfaction. For instance, the "about all they could get" constituted the "satisfied" group, and the "somewhat less than they could get" and "a great deal less than they could get" were combined for the "dissatisfied" group on the basic question for the parent group. This pattern has been followed throughout for the questions. Each summary table following the question being analyzed carries in the final row a figure designated "plus" or "minus" to show the differences in the percentages of the "satisfied" and the "any degree of dissatisfaction." A "plus" denotes satisfaction, while a "minus" indicates dissatisfaction. The analysis of each question follows the same pattern. The percentage of "no opinion" responses is not shown.

1. All things considered, how much (are pupils in your school, is your child, are you) getting out of school work?

Teachers of both groups of students, both groups of parents, and twelfth-grade students themselves expressed the belief that students are getting somewhat less than they could get out of school work. Sixth- and eighth-grade students felt that they were getting about all they could get out of school work.

2. Are there things that you think (the students in your school, your child, you) should be learning right now which are not being taught in your school? (Sixth- and eighth-grade students were not asked to respond to this question.)

A higher percentage of both groups of teachers and the twelfth-grade students were of the opinion that there were "a few — many things" which students should be learning that are not now being taught. Approximately the same percentage of parents of twelfth-grade students expressed this opinion as compared with those who responded "no, none." A higher percentage of parents of sixth- and eighth-grade students responded "no, none," which is an indication of satisfaction.

3. How much of what (the students in your school are, your child is, you are) studying do you think will be useful in everyday living?

A higher percentage of each respondent group was of the opinion that "most to practically everything" that studentx are studying will be useful in everyday living.

4. How free do you feel to use the teaching methods that you believe to be most effective?

An analysis of teacher responses shows that there are very few restrictions placed upon the teaching methods which they choose to use.

5. Are you satisfied or dissatisfied with the teaching methods used in your child's school? In your school? (Sixth- and eighth-grade students were not asked to respond to this question.)

A higher percentage of each group of respondents expressed satisfaction with teaching methods, as compared with those who expressed any degree of dissatisfaction. However, twelfth-grade students were much more critical than either group of parents.

6. How much help do you feel (that you give your students, your child receives, you receive) with school work?

A higher percentage of each group of teachers indicated that they were giving their students all or most of the help needed. Parents and students were in accord with this. However, a somewhat smaller percentage of both groups of parents indicated this belief than did either group of students or teachers.

7. In general (do the students in your school, does your child, do you) have to do too much or too little work in order to "keep up" in school studies?

A higher percentage of each group of respondents indicated that the amount of work required was about right. Of the percentage of those who expressed different opinions, both groups of teachers and the twelfth-grade students felt that too much work was required as compared with too little work, while both groups of parents and sixth- and eighth-grade students felt that too little work was required.

8. On the average (do the students in your school, does your child, do you) spend too much or too little time on homework each day?

A higher percentage of both groups of parents and sixth- and eighth-grade students felt that students spend about the right amount of time on homework each day. A higher percentage of both groups of teachers and the twelfth-grade students indicated that students spend too little time on homework. A comparison of the percentage of responses of "too little" with "too much" shows that a much higher percentage of each respondent group, with the exceptions of the sixth- and eighth-grade students, expressed the belief that too little homework is required.

9. On the average, how much time (do you attempt to require each student to spend, does your child spend, do you spend) on homework each day?

The amount of time that students reported spent on homework each day approximates very closely the time reported by parents. A slightly higher percentage of students spend no time on homework than the percentage of teachers who reported that they require no homework. Except for this time bracket, it appears that students spend more time on homework than is expected by teachers. The amount of time required by teachers and the amount of time spent by students are in inverse proportion to each other. For example, 56 percent of the teachers of the twelfth grade indicated that they attempt to require between one-half and one hour of homework daily. Twenty-eight percent of the twelfth-grade students indicated that they spend this amount of time. On the other hand, twelve percent of the teachers of the twelfth grade indicated that they require between one and one and one-half hours of daily homework. Twenty-six percent of the twelfth-grade students reported that they spend this amount of time. The amount of time spent on homework by the sixth- and eighth-grade students follows this same pattern. However, the amount of time spent by this group more nearly parallels that required by teachers.

10. How many of the students (in the school in which you teach, in the school your child attends, in your school) take part in as many student activities (clubs, parties, plays, athletics, etc.) as you think they should?

A higher percentage of each respondent group, with the exception of the twelfth-grade students, indicated that most or almost all students were participating in as many activities as they should. Only nine percent of the twelfth-grade students expressed this opinion, while 83 percent were of the opinion that one-half or less of the students were participating in as many activities as they should.

11. In general, is the discipline in (the school in which you teach, the school your child attends, your school) too strict or not strict enough?

A higher percentage of each respondent group indicated that discipline was "about right" in the schools. Of those who expressed a different opinion, a higher percentage of each group, with the exception of the sixth- and eighth-grade students, indicated that discipline was not strict enough, as compared with those who indicated that it was too strict.

12. What marks did you earn on your last report card?

It is reasonable to expect the grade distribution in a sample of this size to very closely approach a "normal distribution." However, when the reported grades are plotted, the curve is decidedly skewed to the right. Approximately 20 percent more of each student group received grades of "A" than would normally be expected. The percent of failures reported by sixth- and eighth-grade students is only slightly, if any, more than would be expected in groups of this size. The percent of failures reported by twelfth-grade students is approximately one and one-half percent less than would normally be expected.

13. In general, are you satisfied or dissatisfied with the school in which you teach? your child attends? your school?

A much greater percentage of each group of respondents indicated that they were satisfied with schools, as compared with those who indicated any degree of dissatisfaction.

An analysis of other factors which could affect what students are getting out of school work was made in addition to those shown in Table XIII. These were: equipment and physical facilities needed, custodial services, teachers' acceptance by other faculty members and administrative personnel, time to prepare adequately for teaching, satisfaction with teaching assignment, supervision of student activities and other duties, academic freedom, class size and total work load, and treatment of teachers outside the school by people in the community. In each instance a higher percentage of respondents indicated that they were satisfied than those who indicated any degree of dissatisfaction.

Summary

An analysis of the responses to the basic question, “All things considered, how much (are students in your school, is your child, are you) getting out of school work?”, shows that a higher percentage of each respondent group, with the exception of sixth- and eighth-grade students, believes that students are getting either somewhat less or a great deal less out of school work than there were those who indicated that they were getting about all they could. Teachers of the sixth and eighth grades (61 percent), teachers of the twelfth grade (90 percent), parents of sixth- and eighth-grade students (4 percent), and parents of the twelfth-grade students (15 percent), indicated a higher degree of dissatisfaction in response to the basic question than with any of the factors and/or conditions to which they responded in the related questions. This was also true of the twelfth-grade students (52 percent) with one exception. Seventy-four percent more of the twelfth-graders were of the opinion that one-half or more of the students were not taking part in as many student activities as they should.

In response to the question having to do with overall satisfaction or dissatisfaction with the schools, a far greater percentage of each group of respondents indicated that they were satisfied as compared to those who expressed any degree of dissatisfaction. This same opinion was expressed by sixth- and eighth-grade students and their parents about all the factors and/or conditions to which they were asked to respond in the other related questions. Although the degree of satisfaction varied from question to question, there were only three instances in which one or more of the other respondent groups were not in accord with these opinions. Twelfth-grade students' opinions about activities, reported above, was one of these. The others had to do with opinions about things that students should be learning that are not being taught and the amount of homework done by students.

A higher percentage of teachers of the sixth and eighth grades (48 percent), teachers of the twelfth grade (72 percent), and twelfth-grade students (57 percent), indicated that there were some things that students should be learning which are not being taught than those who responded, “no, none.” Those who expressed this opinion were asked to give examples. Of those complying with this request, more advanced courses in mathematics, such as advanced algebra and trigonometry, another course in science and/or language were the only examples cited by teachers of the twelfth grade and twelfth-grade students. “A foreign language” was the most frequent response by sixth-grade teachers. “Provide opportunities for eighth-grade students to take some of the courses such as Algebra I and general science, which are now taught in the ninth grade,” was the most frequent response of teachers of the eighth grade.

A higher percentage of teachers of the sixth and eighth grades (10 percent), teachers of the twelfth grade (65 percent), and the twelfth-grade students (16 percent), indicated dissatisfaction with the amount of time students are spending on homework than those who indicated that they were spending about the right amount of time. An analysis of the data shown in Table IX shows that the amount of time spent on homework as reported by both parents and students is more than that which teachers indicated they expect students to spend.

Conclusions

With the possible exception of the summary statements having to do with things which students should be learning that are not being taught, and the opinion of twelfth-grade students about the number of students participating in student activities, there is nothing in the data resulting from this inquiry to justify the percentage of parents and the especially high percentage of teachers' and twelfth-grade students' opinions that students are not getting as much out of school work as they should. This is also true of the high percentage of teachers and twelfth-grade students who indicated dissatisfaction with the amount of homework done by students. It can only be assumed that in both instances, dissatisfaction is due to factors other than those studied or that opinions reflect some of the most often repeated criticisms of public education without adequate supporting data.

Teachers expressed a higher degree of satisfaction with those things which tend to reflect their abilities as teachers and less satisfaction with those things which tend to reflect student effort and attitudes.

Parents expressed a higher degree of satisfaction with those things which tend to imply future success for their children and less satisfaction with those things which are immediately observable and could reflect "shortcomings" of the school.

Students expressed a higher degree of satisfaction with those things which tend to reflect their immediate efforts, abilities, and attitudes. They expressed less satisfaction with those things which tend to restrict their activities, as well as intangibles, the results of which are not immediately observable. Dissatisfaction by students with each factor and/or condition studied increased as they progressed from grade to grade.

The opinions of one group of teachers tend to be more nearly like those of the other teacher group than like those of any of the other respondent groups.

The opinions of the parents of sixth- and eighth-grade students tend to be more nearly like those of their children. This is also true of the parents of twelfth-grade students. The opinions of twelfth-grade students and their parents tend to be more nearly like those of the teachers of the twelfth grade; while the opinions of sixth- and eighth-grade students and their parents tend to be more nearly like those of the teachers of these students.

TABLE I

SUMMARY OF RESPONSES TO BASIC QUESTION, QUESTION 1, IN PERCENTAGES

	Teachers		Parents		Students		
	6th, 8th	12th	6th, 8th	12th	6th	8th	12th
About all they could get	17	3	43	38	82	69	21
Somewhat less to great deal less than they could get	78	93	47	53	17	30	73
Satisfied (+) compared with any degree of dissatisfaction (-)	-61	-90	-4	-15	+65	+39	-52

TABLE VIII

SUMMARY OF RESPONSES TO QUESTION 8 IN PERCENTAGES

	Teachers		Parents		Students		
	6th, 8th	12th	6th, 8th	12th	6th	8th	12th
About right	40	15	59	56	76	69	40
Too much	0	1	8	8	16	22	9
Too little	50	79	32	35	6	7	47
Satisfied (+) Dissatisfied (-)	-10	-65	+19	+13	+54	+40	-16

TABLE IX

SUMMARY OF RESPONSES TO QUESTION 9 IN PERCENTAGES

	Teachers		Parents		Students		
	6th, 8th	12th	6th, 8th	12th	6th	8th	12th
None	6	1	3	2	2	3	5
Less than 1/2 hour	14	23	9	8	8	7	11
Between 1/2 and 1 hour	51	56	32	24	24	24	28
Between 1 and 1 1/2 hour	22	12	26	26	17	20	26
Between 1 1/2 and 2 hours	5	5	14	17	16	15	15
More than 2 hours	0	0	0	0	11	13	10

TABLE X

SUMMARY OF RESPONSES TO QUESTION 10 IN PERCENTAGES

	Teachers		Parents		Students		
	6th, 8th	12th	6th, 8th	12th	6th	8th	12th
Most or almost all	53	53	68	72	61	58	9
About one-half or less	27	43	20	18	36	41	83
Satisfied (+) Dissatisfied (-)	+26	+10	+48	+54	+25	+17	-74

TABLE XI

SUMMARY OF RESPONSES TO QUESTION 11 IN PERCENTAGES

	Teachers		Parents		Students		
	6th, 8th	12th	6th, 8th	12th	6th	8th	12th
About right	72	53	70	60	82	78	60
Too strict	0	0	3	4	12	14	15
Not strict enough	23	43	26	30	4	6	19
Satisfied (+) Dissatisfied (-)	+49	+10	+41	+26	+66	+58	+26

TABLE XII

SUMMARY OF RESPONSES TO QUESTION 12 IN PERCENTAGES

	Mostly D's or lower	Mostly C's and D's	Mostly C's	Mostly B's and C's	Mostly B's	Mostly A's and B's	Mostly A's
6th Grade	5	8	5	20	8	30	23
8th Grade	7	9	6	19	10	25	23
12th Grade	2	5	6	22	15	28	22

TABLE XIII

SUMMARY OF RESPONSES TO QUESTION 13 IN PERCENTAGES

	Teachers		Parents		Students		
	6th, 8th	12th	6th, 8th	12th	6th	8th	12th
Satisfied	85	82	80	80	83	78	70
Dissatisfied	13	16	17	16	16	22	29
Satisfied (+) Dissatisfied (-)	+72	+66	+63	+64	+67	+56	+41

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 3, November 1962)

Development of and Reasons for Vocational Choices Of Adolescents Through the High-School Years

MARVIN POWELL
Northern Illinois University
VIOLA BLOOM
Elmhurst College

THE INTENT of this study is to investigate the development of the vocational plans of adolescents, to determine the objectivity of their choices, to discover the motivational forces operating within vocational outlook, and to locate any pattern of problems that are present in this area of adjustment.

Review of Previous Literature

A review of the previous literature reveals a number of studies relating in one way or the other to the present study. Various investigators — Powell (8), Super (11), Kuhlen (3), Shaffer (9) — have been concerned with adolescents' vocational choices, particularly in relationship to general adjustment. Studies of the reasons for choice or certain factors related to choice have also been published — Ezell and Tate (1), Norton (6). Stephenson (10) found a lack of intelligent planning in ninth graders. Moser (5) studied influences of cultured factors upon selection of vocational preferences. Many other studies — Ginsberg (2), Levine and Wallen (4), Peters (7), etc. —, have presented a number of aspects of vocational preferences.

The present study in some areas repeats previous research, but in a modern day — (1959)— frame of reference. Further, the present study indicates grade trends from tenth through twelfth grades.

Procedure

A questionnaire was devised so that the student could answer each question with complete freedom of response. The questionnaire was administered to approximately one thousand students of varied socio-economic levels which included students in industrial areas and those in upper-class suburban communities. Since the study was intended as a pilot study for further research, no effort was made to control the sampling bias. The results of the questionnaire were then tabulated and conclusions derived. The attempt was made to arrive at significant conclusions which were to serve as the guide for further, more intensive research in the area of psychological adjustment to vocational outlook.

The Sample

The questionnaire was administered to 929 high school students in the tenth, eleventh and twelfth grades in a small metropolitan school district with the distribution as shown in Table I. The ages of the students ranged from 14 through 19 years. The education of the parents was used to give an estimate of the socio-economic status of the sampling. Table II shows that there are proportionately more parents who have attended college than in the general population.

Results

What occupation would you like to enter assuming that you have financial resources, ability and freedom of choice?

The occupations were classified according to the descriptions given in the Dictionary of Occupational Titles. The analysis of these responses is shown in Table III. The specific occupations selected by the boys were: 1) engineering — 23.3 percent; 2) medicine and automotive mechanics —

each 5.6 percent; 3) research in physics and chemistry — 4.1 percent; 4) aviation and "the business world" — each 3.4 percent. The phrase, "to enter the business world," was often given by boys with little indication of its meaning other than the idea of a white collar position.

The choices of occupation as given by the boys were dispersed over 50 different fields with 33 of the named occupations being ranked on the professional level. About 6.6 percent of the boys selected various fields of science as biology, archaeology, astronomy and geology. The occupations which were given on the skilled labor level included those of electrician, carpenter, television serviceman and automotive mechanic.

The vocations indicated as most preferred by girls were: 1) office work — 21.6 percent; 2) teaching — 15.0 percent; 3) nursing — 13.7 percent. These three were chosen significantly more often than other occupations. The fourth in rank was that of airline hostess with 4.1 percent of the girls selecting it as the preferred position. There were about 35 different vocations named by the girls; 33 of these were of professional status. Some of the other vocations named were that of journalism, modeling, dramatics, fashion designer, interior decorator , physical therapist, lawyer, librarian and social worker.

What occupation do you really expect to enter?

Reference to Table III will show the analysis of the vocational level which the students are planning to enter. Although the general trend is the same, with the emphasis on the professional occupations, there are significant differences between the desired and the expected vocation. Significantly fewer boys expect to enter occupations categorized as professional and skilled. There are significantly more who expect to enter the clerical-sales classification, the semi-skilled trades and the armed services. Approximately nine percent frankly admitted that they had no idea of their vocational future. The specific occupations listed as those which they plan to enter are: 1) engineering — 18.6 percent; 2) automotive mechanic — 5.4 percent; 3) to enter the business world — 4.7 percent; 4) salesman — 4.5 percent; and semi-skilled labor as factory work — 4.1 percent. About 7.7 percent of the boys expected to enlist in the armed services upon completion of high school in comparison to 3.2 percent who had stated that they preferred to enlist.

In a scrutiny of the vocations which the girls expect to enter, there is also a repetition of the same curve as the occupations they preferred. However, there were significant differences noted. Significantly fewer girls planned to enter professional vocations while significantly more were preparing to enter clerical-sales level. It is interesting to observe that there is also a significant increase in those girls who expected to be a housewife. Seven percent of the girls did not know what they would do upon graduation.

The vocations which the girls are preparing to enter are: 1) office and clerical work — 29.8 percent; 2) teaching — 16.9 percent; 3) nursing-12.8 percent; and 4) housewife — 5.5 percent.

Do adolescents expect to enter the vocations they prefer? What are the reasons for not entering the desired vocations?

Approximately 63 percent of boys and girls expect to enter the vocation that they prefer. The proportion of those actually preferring the vocation that they expect to enter increases with age. In the tenth grade, about 59 percent seem to plan to enter the vocation they prefer; in the eleventh grade the proportion rises to 63 percent; and, in the twelfth grade the proportion increases to 67 percent.

As shown in Table IV, the reasons given by the boys for the discrepancy between desired occupations and the occupation they expect to enter are: 1) an interest in the work they plan to enter — 18.4 percent; 2) undecided as to the type of work they will enter — 15.8 percent; 3) the vocation they are to enter is the choice of the parents — 11.1 percent; 4) the decision to enlist in the armed services immediately — 10.5 percent; 5) inadequate finances for the preparation needed — 7.9 percent; 6) having ability for the work they are to enter — 5.8 percent. Other reasons named are, in order of frequency, the opportunity presented by the chosen occupation, high financial returns in the vocation they plan to enter, plan to enter a related field of work, need to have work preparatory for the preferred position, lack of ability for the preferred job and the desire for the security found in the expected vocation.

Girls name the same general reasons for not entering the vocation which they prefer but place them in a different order of importance. This order is: 1) inadequate finances to prepare for the desired position — 15.4 percent; 2) plans to marry and the necessity of entering preparatory work for the vocation — each 12.8 percent; 3) interest in the occupation they plan to enter — 10.3 percent; and 4) having the ability to perform the work they are to enter.

Why do you prefer to enter this vocation?

In an attempt to arrive at some of the motivational forces, the students were requested to give the reason that they preferred the certain vocation. The responses are summarized in Table V.

There are a few significant trends which deserve attention. With regard to the responses given by the boys, the selection of a job because of an interest in people shows an increase from

1.3 percent of the total given for liking a job in the tenth grade to 5.7 percent in the twelfth. This is not a significant trend but does indicate an increasing social consciousness. Interest in people is also reflected in the increase in the desire for a job in which the individual may be of service to others. There is a significant trend at the .01 level of confidence between the eleventh and twelfth grades in the frequency that boys gave service to others as a reason for desiring a particular job. During the eleventh grade, there is a significant increase in two factors. The desire for adventure and travel also shows a sudden increase in frequency in the eleventh grade, although this is not quite significant at the .05 level of confidence.

The greater social awareness of girls is seen in the jobs they prefer as well as in the motives they give for selection of a vocation. This interest in working in a social-centered atmosphere appears in the eleventh grade for there is a significant increase in the frequency of recognition of this motive during this stage of development (.01 level).

The desire for financial security shows a significant decrease with age — at the .05 level of confidence between the tenth and eleventh grades and at the .01 level between the tenth and the twelfth grades. There is also an indication that the ability to do the work assumes greater importance with age, for the girls give it as a reason for choosing a certain job 2.1 percent of the time in the tenth grade, 5.3 percent in the eleventh grade, and 9.3 percent in the twelfth grade. Girls also mention the desire for a certain vocation because of its adaptability to married life. Nursing and teaching are occupations which are chosen because, "you can work when you have the time," or, "the vacations are the same as children's school vacations so you can be with the children."

In a comparison of the differences of motives given by boys and girls, it is found that boys claim an interest in the work to be more important than do the girls. This difference is significant at the .01 level. Girls place the motive of an interest in people and a desire to be of service to others significantly more than do boys; this is significant at the .01 level of confidence. Girls are also significantly more influenced by the desire for security. Boys acknowledge security as a motive force about 5.3 percent of the time whereas girls give it 17.4 percent of the time.

What is your ultimate goal with regard to vocational objectives?

In an attempt to determine whether the students had long range life aims, they were asked to specify their vocational goals. The boys responded as follows: 1) to enter a specific vocation — 21.4 percent; 2) to have their own business — 13.1 percent; 3) to become a personal success in their position — 11.9 percent; 4) the desire for financial success and security — 11.1 percent; 5) to have a good measure of social prestige — 4.9 percent; and 6) to do the job well. Other reasons which a few of the high school youth mentioned were: to be of service to others; to be creative in some work that is accomplished; to finish college; to be content in the work, and to achieve a degree of happiness.

The desire for financial gains were expressed as "wanting enough money to buy anything I want," or "having enough money to live comfortably." Some stressed the acquisition of more money than they would need while others emphasized economic security. Fifteen percent of the boys did not specify any vocational objective. The responses of the girls with regard to an ultimate objective were: 1) the desire to enter a specific vocation — 37.6 percent; 2) the desire to do the job well — 9.3 percent; 3) the desire for personal success and advancement — 8.1 percent; 4) the desire to marry and have a home and children — 7.6 percent; 5) to be of service and to help others — 5.7 percent. Other responses, in order of proportionate frequency, include finishing college, enjoying the work, having financial success, needing security and gaining social prestige. Eighteen percent gave no answer.

Who has influenced you in your selection of a vocation?

Reference to Table VI will show that the majority of high school students do not acknowledge that any one has influenced them in the selection of a vocation and that it is their own decision.

Have you had vocational counseling? By whom? Did it help in the selection of a vocation?

In response to these questions, 44.6 percent of the boys and 52.2 percent of the girls claimed they had received vocational guidance. In reporting the source of the guidance, 27.5 percent said it originated with a teacher, 21.1 percent attributed it to the school counselor and 17.0 percent credited their parents with guiding them. Only 5.5 percent stated they had vocational counseling from a professional counselor or from a vocational conference. The major emphasis was on the incidental counseling received at school. Others, such as a minister, a friend, or business associate played a minor role.

In the evaluation of the counseling received, 45.7 percent stated that it was of positive help to them; 12.5 percent reported the counseling was of little help; and 40.2 percent stated that it did not aid them at all in the selection of a vocation.

What factors do you consider of importance in a vocation?

Table VI gives an analysis of the responses to this inquiry. There are some trends which are worthy of comment. The importance placed on security decreases significantly for the boys from the tenth through the twelfth year of school; working conditions, per se, are of more consideration to the tenth-grade male students than to the twelfth-grade students; knowledge of the job receives significantly more attention from boys in the eleventh grade than in the tenth or twelfth; a true interest in the work receives significantly greater emphasis with twelfth-grade boys than with others.

Emphasis on the acquisition of a good salary receives significantly more attention from tenth-grade girls than girls in the other years of school; there is added emphasis on the ability to do the job as girls grow older; with increasing age there is more importance assigned to the possibilities of personal advancement; also with increasing age there is a significant decrease in the stress on the ability to get along with people and, with the increase of age, girls tend to place more emphasis on the degree of happiness and enjoyment which they can receive from a vocation.

In comparing the differences between what the boys and the girls consider as important, it will be noticed that boys place significantly more emphasis on the financial aspects of a vocation.

Concluding Comments

Although there is a bias in the sampling toward the upper end of the educational scale, the 50 percent of students who expect to enter professional vocations is higher than what is to be expected. There are significant differences between planned and preferred vocations which show a reality orientation of high-school youth. This is an adjustment of objectives with personal resources. There is also a lack of reality orientation depicted by the 20 percent who plan to enter skilled, semi-skilled or unskilled vocations although 63 percent of the parents had high school education or less. Since most students planned to enter the vocation they preferred, it may be proposed that they did not feel free to speculate concerning the occupation desired or that they did not respond with objective rationality to the problem of planned vocation.

Insight is gained into the perception of adolescents in the designation of motives. Most responses centered on an interest in the work, but the question remains of the why of the interest factor. Why are students attracted to certain vocations? What motives are in operation? In the instance where boys select engineering, the effects of wide advertising are apparent. Emphasis on engineering is in accordance with the cultural role of the male, a force toward mechanical things. Girls are more in conformity with accepted cultural roles depicted by the narrow range of vocations expressed. Nursing, teaching and secretarial work are accepted positions for women. The motives given are prosaic and show that there is little insight or personal awareness. Cultural factors, the desire for prestige or an avarice for material possessions are not recognized as important.

This stage of development is characterized by the increased social awareness and is readily observed in the responses of girls. It is more subtle in answers given by boys. The urge for independence is elicited in their desire for travel and adventure. Independence, probably the most potent motivational force, can also be observed by their reticence to credit others with having an influence over their vocational selection. This seems to be a prime indication of adulthood to the adolescent.

The desire for social prestige may be observed in the need of girls to select a vocation which is valued by society. Approval is gained through service to others. Engineering offers social approval for the male. Our present technical culture esteems this above all others. Is interest, as given by boys and girls, a reason for selecting a vocation, including the desire for social approval? Does interest in a vocation represent a means of winning social prestige?

The problem areas which seem to arise from this study are:

1. Youth are frustrated in the intelligent selection of a vocation because of the lack of knowledge of vocational fields. Conflicts arise when the individual is not aware of the breadth of opportunities in each vocational field and thus the youth become inhibited in entering an occupation of his choice. There is definite need for more vocational orientation in junior and senior high school.

2. Adolescents are not aware of the motives for selecting vocations. They can not recognize the inconsistency of motive and ability. Further research is recommended in this area. What motives and influences are actually operating besides the obvious ones? How can motives be recognized and redirected into a more suitable vocational outlook? Are certain personality types attracted to similar vocations? Can motives be altered through effective guidance and unfavorable influences overcome?

3. In spite of emphasis on guidance in school, adolescents have had little vocational guidance and seldom is it effective. Is vocational guidance a function of the school? Are counselors trained to give effective guidance? Are school programs designed to provide real counseling? Some attention should be given these matters. If vocation guidance is accepted as a function of our present school system, measures ought to be taken

to ascertain the value and standard of the counseling.

4. The lack of effective guidance coupled with an inability to recognize motives inhibits the students in establishing long-range vocational objectives. The advanced degree of our culture demands that many be willing to postpone goal fulfillment through the need of many years of education. The goals as specified by the students are not mature in outlook nor do they represent an understanding of adult culture. The very obscurity of the responses demands additional study to discover the degree of knowledge of the students in regard to the establishment of vocational objectives. This should direct the path of counseling to remedial procedures.

REFERENCES

1. Ezell, L. B., and Tate, H. H. "High School Students Look to the Future," Journal of Educational Research, XLIX (1955), pp. 217-222.
2. Ginsberg, Eli, et al. "The Problem of Occupational Choice," American Journal of Orthopsychiatry, XX(1950), pp. 166-201.
3. Kuhlen, R. G. The Psychology of Adolescent Development. (New York: Harper and Bros., 1952), pp. 244-245.
4. Levine, Phyllis, and Wallen, Richard. "Adolescent Vocational Interests and Later Occupations," Journal of Applied Psychology, XXXVIII (1954), pp. 428-431.
5. Moser, William. "The Influence of Certain Cultural Factors upon the Selection of Vocational Preferences by High School Students," Journal of Educational Research, XLV (1952).
6. Norton, J. L. "General Motives and Influences in Vocational Development," Journal of Genetic Psychology, LXXXII (1953), pp. 263-278.
7. Peters, E. F. "Factors Which Contribute to Youth's Vocational Choices," Journal of Applied Psychology, XXV (1941), pp. 428-430.
8. Powell, Marvin. "Age and Sex Differences in Degree of Conflict within Certain Areas of Psychological Adjustment," Psychological Monographs, LXIX, No. 2 (1955).
9. Shaffer, L. F. The Psychology of Adjustment (Boston: Houghton-Mifflin Co., 1936).
10. Stephenson, R. M., Jr. "Occupational Aspirations and Plans of 443 Ninth Graders," Journal of Educational Research, XLIX (1955).
11. Super, Donald. "Occupational Level and Job Satisfaction," Journal of Applied Psychology, XXIII (1939), p. 550.

TABLE I

THE DISTRIBUTION OF STUDENTS

Grade	10	11	12	Total
Boys	208	174	176	558
Girls	175	69	122	366
Total	383	243	298	924

TABLE II

THE PERCENTAGE OF PARENTS COMPLETING SPECIFIED NUMBER OF YEARS IN SCHOOL

Years Completed in School	Less than 8	8-11	12	13 or More	Not Known
Mothers	3.0	21.2	42.5	25.5	7.7
Fathers	4.9	21.8	33.9	30.8	8.7
Total	3.95	21.48	38.20	29.81	8.1

TABLE III

THE PERCENTAGE OF STUDENTS WHO PREFER AND PLAN TO ENTER CERTAIN VOCATIONAL LEVELS

	Boys	Boys	Girls	Girls
	Occupation Preferred	Occupation Planned	Occupation Preferred	Occupation Planned
Professional	64.9*	49.3	69.7	53.0*
Clerical - Sales	6.6	9.9**	23.2	30.9**
Service	1.3	1.3	1.6	1.4
Agriculture	5.4	4.7	.6	.6
Skilled	12.5*	7.9		
Semi-Skilled	3.1	9.5**		
Unskilled	.8	.9		
Armed Services	3.2	7.7**	1.4	1.4
Housewife			1.6	5.5*
No Response	2.0	9.0*	1.9	7.4*

* Significant difference at .01 level between Preferred and Planned Vocation.
** Significant difference at .05 level between Preferred and Planned Vocation.

TABLE IV

REASONS GIVEN FOR NOT ENTERING THE PREFERRED OCCUPATION
(PERCENTAGES)

	Boys	Girls
Inadequate finances	7.90	15.39
Ability for chosen work	5.79	9.40
Lack of ability for chosen work	2.63	5.98
Choice of parents	11.05	2.56
Preparatory work for job	3.16	12.82
Opportunity in chosen work	4.74	5.98
Like chosen work	18.42	10.26
Marriage	0.00	12.82
High Pay	4.74	.86
Security	2.10	.86
Undecided	15.79	6.84
No reason	6.32	6.84
Armed services	10.53	0.00
Miscellaneous	3.16	6.84
Related field	3.68	2.56

TABLE V

THE PERCENTAGE OF TOTAL REASONS GIVEN AS A BASIS FOR THE SELECTION OF CERTAIN VOCATIONAL AREAS

Motivating Factor	Boys	Girls
Interest in work	36.2	27.4
Ability to do the work	6.1	4.3
Interest in people	3.4	15.2
Personal advancement	7.0	3.9
Offers security	11.9	6.0
Service to others	5.3	17.4
Adventure and travel	4.3	5.1
Creative idealism	5.3	4.0
Plan to marry	.1	5.8
Armed services	2.7	.6
Previous experience	3.6	2.6
Variety within the work	1.5	.6
Need education	.4	1.4
Religious	.9	.6
Social prestige	.6	.7
Same work as father	3.3	.5
Parental influence	1.6	1.0
Miscellaneous	2.5	1.9
No reason given	5.2	2.9

TABLE VI

PERSONAL INFLUENCES ON HIGH SCHOOL STUDENTS IN THEIR CHOICE OF VOCATIONS (IN PERCENTAGES)

	Boys	Girls
Mother	1.79	8.47
Father	15.23	2.73
Both parents	4.66	7.92
Other relative	7.71	10.11
Friend	2.15	4.64
Doctor	.72	.27
Minister	.90	.82
Teacher	5.20	6.56
Vocational guidance	1.08	1.91
Professional acquaintance	6.09	3.01
No one	42.65	40.92
Work experience	1.25	.82
No answer	6.27	6.01
Miscellaneous	4.30	5.74

TABLE VII

FACTORS WHICH HIGH SCHOOL STUDENTS CONSIDER AS IMPORTANT IN A VOCATION (IN PERCENTAGES OF FREQUENCY)

Designated Factors	Boys	Girls
Financial reward	13.43	8.87
Need of security	8.37	6.37
Variety of duties	.40	.34
Good working conditions	5.77	4.55
Ability to do the job	6.95	8.99
Knowledge of the job	8.93	3.30
Opportunity for advancement	6.56	5.80
Social prestige	1.66	1.14
Friendliness	3.00	5.23
Travel and excitement	.47	.11
Need of education	2.92	2.28
To achieve a goal	1.34	3.19
To benefit others	.95	3.53
True interest in work	8.53	8.65
Satisfaction	3.32	4.21
Happiness	2.69	4.55
Enjoyment	13.43	15.81
Miscellaneous personality characteristics	11.30	13.08

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 3, November 1962)

The Relation Between Junior High-School Students' Problems and Parental Attitudes Toward Child Rearing and Family Life

MICHAEL ZUNICH
Texas Woman's University

A REVIEW of the literature indicates that there have been numerous studies concerned with the measurement of parental attitudes and problems of students (1, 2, 5, 6, 7). However, comparatively little research has been done in which an attempt has been made to relate the influence of parental attitudes to the problems of students. Frequently, studies of parental attitudes have been concerned with younger children, and/or investigators have tried to relate some variable such as sex, age, and personality to the problems of students. Yet it is commonly assumed by many educators and psychologists that : a) the practices which parents use in their treatment of their children serve as originative forces on the child's behavior (e.g., the techniques of control and punishment, values they impose, and the incentives they offer), and b) that there is a close relation between attitudes of parents toward their children and the adequacy of the children's social and emotional adjustment.

This study was designed to test the hypothesis that parental attitudes toward child rearing and family life are significantly related to problems of junior high-school students.

Sample and Procedure

The subjects consisted of 40 junior high-school students and their parents. The students were randomly selected from the junior high schools in Denton, Texas. Of these students, 20 were girls and 20 were boys ranging in age from 12 to 14. The parents of these students were: a) middle-class, b) white, c) between the ages of 32 - 50, and d) from intact families with one or more children. The McGuire-White (3) Index of Social Status (Short Form), was utilized in obtaining a measurement of socio-economic status.

In order to obtain measures of problems, students were given the Mooney Problem Check List (MPCL), Junior High School Form (1950 Revision), developed by Mooney and Gordon (4). The MPCL consists of 210 problems divided into seven areas with 30 items in each area. The seven areas are: 1) Health and Physical Development, 2) School, 3) Home and Family, 4) Money, Work, the Future, 5) Boy and Girl Relationship, 6) Relations to People in General, and 7) Self-Centered Concerns.

The Parental Attitude Research Instrument (PARI), developed by Schaefer and Bell (8), was utilized in order to obtain measures of parental attitudes toward child rearing and family life.

The Mooney Problem Check List was administered to the students by the investigator during regular class sessions. The students were instructed to underline the problems troubling them most frequently. After the students completed the MPCL, their parents were given the Parental Attitude Research Instrument. The PARI was administered to the parents by means of a structured interview. The instrument was given separately to each parent on two different occasions.

Results and Conclusions

Of the 210 problems, girls evidenced a higher frequency of problems. Girls were more concerned with problems of School, Home and Family, Money, Work, the Future, Boy and Girl Relationship, and Relations to People in General. However, boys had a higher frequency of problems dealing with Health and Physical Development, and Self-Centered Concerns. The highest number of problems for the girls appeared in the area of Relations to People in General, and the lowest in the area of Self-Centered Concerns. Boys, on

TABLE I

SIGNIFICANT CORRELATION VALUES BETWEEN MATERNAL ATTITUDE SCORES AND GIRLS' RESPONSES TO PROBLEMS

	Mooney Problem Check List Areas						
Attitude Subscales	Health, Physical Development	School	Home and Family	Money, Work, the Future	Boy, Girl Relationship	Relations to People in General	Self-Centered Concerns
Acceleration of Development	.62**						
Ascendancy of Mother	.45*						
Avoidance of Communication	.65**			.56**	.56**	.51**	.60**
Breaking the Will	.47*			.49*			
Comradeship and Sharing			.43*	.57**	.49*		
Fear of Harming Baby	.63**			.56**	.56**	.59**	.66**
Intrusiveness	.45*						
Irritability	.49*						
Marital Conflict				.43*			
Martyrdom	.43*			.54*			.45*
Rejection of Homemaking	.46*			.44*			
Seclusion of Mother				.43*			
Suppression of Aggression	.43*			.54*			
Suppression of Sex				.46*			

* Significant at .05 level
** Significant at .01 level

TABLE II

SIGNIFICANT CORRELATION VALUES BETWEEN PATERNAL ATTITUDE SCORES AND GIRLS' RESPONSES TO PROBLEMS

Attitude Subscales	Mooney Problem Check List Areas						
	Health, Physical Development	School	Home and Family	Money, Work, the Future	Boy, Girl Relationship	Relations to People in General	Self-Centered Concerns
Ascendancy of Mother	.56**	.43*		.43*			.43*
Avoidance of Communication	.59**			.49*	.45*		.50*
Breaking the Will				.52*			
Comradeship and Sharing		.47*		.69**	.46*		.57**
Dependency of Mother	.48*	.47*		.54*	.54*	.56**	.62**
Equalitarianism	.44*						
Excluding Outside Influence				.46*			
Fear of Harming Baby	.53*			.59**			.50*
Fostering Dependency		.46*		.51*			
Inconsiderateness of Husband			.45*				
Intrusiveness	.52*						
Marital Conflict				.49*			
Rejection of Homemaking		.43*		.66**			
Suppression of Aggression	.52*			.73**	.43*	.45*	.53*
Suppression of Sex				.71**			.44*

* Significant at .05 level
** Significant at .01 level

TABLE III

SIGNIFICANT CORRELATION VALUES BETWEEN MATERNAL ATTITUDE SCORES AND BOYS' RESPONSES TO PROBLEMS

Attitude Subscales	Mooney Problem Check List Areas						
	Health, Physical Development	School	Home and Family	Money, Work, the Future	Boy, Girl Relation-ship	Relations to People in General	Self-Centered Concerns
Acceleration of Development	.54*	.68**		.50*	.55**	.60**	.59**
Avoidance of Communication		.47*					
Encouraging Verbalization			.43*				
Excluding Outside Influence		.50*					
Irritability		.50*					
Suppression of Sex						.46*	

* Significant at .05 level
** Significant at .01 level

TABLE IV

SIGNIFICANT CORRELATION VALUES BETWEEN PATERNAL ATTITUDE SCORES AND BOYS' RESPONSES TO PROBLEMS

Attitude Subscales	Mooney Problem Check List Areas						
	Health, Physical Development	School	Home and Family	Money, Work, the Future	Boy, Girl Relation-ship	Relations to People in General	Self-Centered Concerns
Acceleration of Development	.49*		.46*	.46*		.55**	
Avoidance of Communication			.50*	.46*			
Breaking the Will	.48*				.48*	.51*	
Encouraging Verbalization		.44*			.44*		
Irritability	.45*						
Martyrdom						.46*	
Rejection of Homemaking					.43*		

* Significant at .05 level
** Significant at .01 level

the other hand, had the highest frequency of problems concerning Self-Centered Concerns, and the lowest number in the area of Boy and Girl Relationship.

Relation between Parental Attitudes and Students' Problems

The significant correlation values between parental attitudes and students' problems are presented in Tables I to IV. Comparisons were made by means of Pearson product-moment correlation coefficients computed between frequencies in the MPCL areas and the attitude subscales. Of the 644 comparisons, 92 evidenced significant relations at the .05 level or beyond.

Sixty-seven significant correlation values were observed between parental attitude scores and girls' responses to problems (see Tables I and II). Of these relations, fathers and daughters evidenced 38 relations , while mothers and daughters evidenced 29.

Although many of the significant relations were the same for maternal and parental attitude scores and girls' problems, some variations were noted. For example, the most frequent relation, for both parents, was observed between the two MPCL areas of Health and Physical Development and Money, Work, the Future, and the PARI subscales. However, in the two problem areas of School and Self-Centered Concerns, three relations only appeared between mothers and daughters, whereas twelve relations were observed between fathers and daughters.

Twenty-five significant relations were observed between maternal and paternal attitude scores and boys' responses to problems (see Tables III and IV). Of these relations, ten significant correlation values were observed between the attitude subscale Acceleration of Development and the seven problem areas. Mothers and sons evidenced eleven relations while fathers and sons fourteen. It is interesting to note that there were 42 less relations between parental attitude scores and boys' problems than between parental scores and girls' problems.

It will be noted from an examination of Tables I to IV that some of the significant relations have no logical connection. For example, the PARI subscale Suppression of Sex and the MPCL area Money, Work, the Future, have no logical meaning. These relations may be due to chance, i.e., one might expect to observe values, by chance, which attained significance at the .05 level approximately five times out of 100 when no real relation exists. Also, the lack of logical connection may be due to the MPCL areas since they were not designed specifically to yield measures of phenomena assessed by the PARI.

The purpose of this investigation was to test the hypothesis that parental attitudes toward child rearing and family life are related to problems of junior high-school students. The implicit prediction in this hypothesis was clearly borne out that students' problems are related to parental attitudes. Of the 644 relations examined between parental attitude scores and students' problems, 92 were significant at the .05 level or beyond. Among the relations examined for fathers' attitude scores and students' problems, 52 were significant, whereas only 40 were significant when the same number of relations were examined for mothers and students. The most revealing data were found between parental attitude scores and girls' responses to problems. Of the relations observed between parents and their daughters, 67 were significant while only 25 relations were noted between parents and sons.

REFERENCES

1. Anderson, John E. "Parents' Attitudes on Child Behavior; A Report of Three Studies," *Child Development*, XVII (1946), pp. 91-97.
2. Burchinal, L. R., and Hawkes, Bruce R. "Marriage Adjustment; Personality Characteristics of Parents and the Personality Adjustment of Their Children," *Marriage and Family Living*, XIX (1957), pp. 366-72.
3. McGuire, C., and White, G. *The Measurement of Social Status*, Research Paper in *Human Development*, No. 3 (Revised), (University of Texas, Austin, Texas, 1955).
4. Mooney, Ross, and Gordon, Leonard. *Mooney Problem Check List*, 1950 Revision (Ohio State University, 1950).
5. Radke, Marian J. "Relation of Parental Authority to Children's Behavior and Attitudes," *University of Minnesota Institute of Child Welfare Monographs* (University of Minnesota, No. 22, 1946).
6. Read, Katherine H. "Parents' Expressed Attitudes and Children's Behavior," *Journal of Consulting Psychology*, IX (1945), pp. 95-100.
7. Rouman, Jack. "School Child's Problems as Related to Parental Factors," *Journal of Educational Research*, L (1956), pp. 105-12.
8. Schaefer, Earl S., and Bell, Richard Q. "Development of a Parental Research Instrument," *Child Development*, XXIV (1958), pp. 339-61.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 3, November 1962)

An Experiment in Teaching Elementary School Methods via Closed Circuit Television*

HERBERT P. RUMFORD
Kansas State College of Pittsburg

A RAPIDLY GROWING student population has caused American educators to consider the feasibility of using various media of mass communication for some phases of instructional programs. One area in which increased enrollment has created a serious problem is in teacher education, where students are expected to observe the techniques of classroom procedure. It has been proposed that prospective teachers might observe classroom procedure via television as a practical solution.

Several armed services investigations by Kanner (5), Manley (8), Rock, et al. (10), and many college and university studies such as Carpenter and Greenhill (1, 2), Erickson (3), Macomber (7), and The Minnesotan (9), have examined the effectiveness of TV for teaching the content of lecture-type courses. The present investigator knows of no published study which reports the effectiveness of television as a means of observing teaching methods.

Selected for study in this experiment was Elementary School Methods, a sophomore course, which met for two hours each day, Monday through Friday, throughout the fall semester of 1957. This two-hour period was used in two ways: from 8:00 to 8:50, each morning, the instructor lectured on methods of teaching; and from 9:00 to 9:50 the students observed classroom teaching in the elementary laboratory school which is operated by the college. The course was divided into six units which included: Arithmetic, Kindergarten Methods, Primary Reading, Intermediate Reading, Language Arts, and Social Studies.

Method

Subjects. The effectiveness of total teaching via television was measured by comparing the achievement of matched groups of students under controlled conditions. An experimental group and a control group were formed by matching fifteen pairs of individuals, who had had no methods course experience or teaching experience, on the basis of intelligence test scores, age, sex, and scores made on the Minnesota Teacher Attitude Inventory. In the matching procedure, precedence was given to intelligence scores.

Design and Materials. Lectures originated in a studio-classroom of the Audio-Visual Center with the control group present. The classroom demonstrations originated in a room set apart for television programs in the laboratory school with the control group present. This room was furnished with chairs and tables best suited for the age level of the children who constituted the class for that day. The television programs, both lectures and demonstrations, were viewed by the experimental group in a receiving room in the main college classroom building.

The principal of the elementary laboratory school lectured throughout the semester and served as the instructor for the course. He was responsible for planning the demonstrations of elementary methods and assisted the camerman in selecting the video portion to be televised.

The classes used for the demonstrations included those from kindergarten through the sixth grade. All seven instructors held master's degrees in education and all of them served as supervisors of college classes in student teaching. These demonstrating instructors assisted the principal in compiling objective questions, sampling what seemed to them important in teaching methods, and scoring the daily observation papers written on the day they demonstrated teaching methods for the course.

---

* This work was done as partial fulfillment of requirements for an Ed.D. degree at the University of Colorado, 1960.

In both the Audio-Visual Center and the laboratory school, the television equipment was low-cost vidicon industrial-type manufactured by the Dage Television Division of Thompson Products, Incorporated. Two cameras were installed in the Audio-Visual Center and three in the laboratory school. The receiving room was equipped with three standard RCA receiving sets with 24-inch screens. The signal was sent from the originating room to the receiving sets over coaxial cable. The audio signal was carried over a transmission line and fed through a separate amplifier.

Responsible for the technical operation of the television equipment was the Audio-Visual Director of the college. He was assisted by a full-time electronics technician and several part-time helpers.

Procedure. The effectiveness of television in affording opportunities to make significant observations was measured by the statistical analysis of differences between the scores of the study groups obtained from daily observation papers and from objective tests over observations.

The effectiveness of television as a means of teaching the content of the lectures on elementary school methods was measured by the statistical analysis of differences between the scores of the study groups, obtained from multiple-choice, teacher-made tests.

At the beginning of the semester, before any course work was started, and again at the end of the semester, the Minnesota Teacher Attitude Inventory and the Minnesota Multiphasic Personality Inventory were administered to all the students enrolled in the methods course. Scores from the MTAI and selected scales from the MMPI were used as a measure of personality characteristics.

Changes in personality that might be attributed to experience related to instruction via television were assessed by the statistical analysis of pretest-post-test difference scores of the study groups for selected scales of the Minnesota Multiphasic Personality Inventory and the Minnesota Teacher Attitude Inventory.

The t-test was used to test the significance of differences in related means and variances, and the Wilcoxon matched-pairs signed-ranks test was used as a non-parametric test for the significance of differences when the groups were considered separately by ability levels (high seven and low seven in intelligence test scores).

Results

Achievement. Table I shows means and t values for the differences between means when the unit tests over the lectures, the unit tests over observations, and observation points were summed. The differences between mean scores for the tests on the lectures and tests over observations proved non-significant, but the television group earned more observation credits than were earned by the group in the originating classroom. The difference was significant at the .05 level.

Table II shows means and t values for the difference between means for the final test scores. All differences between intact groups proved non-significant. However, when the scores from the test on observations were examined according to ability levels (high seven and low seven), it appeared that the high seven in the originating room exceeded the high seven in the television room. Wilcoxon T equalled 1 which is significant at the .05 level.

Table III shows means and t values for the differences between means for the final test given, as a test of retention, four months after the end of the semester. Differences for intact groups were non-significant. When the scores from the test on observations were examined according to ability levels, it appeared that the high seven in the originating room exceeded the high seven in the television room. Wilcoxon T equalled 0 which is significant at the .05 level. This is the same indication noted at the end of the study semester.

Personality Variables. Table IV shows the mean gain in MTAI scores for both groups and the t values of the differences in means. Whereas both groups made significant gains in MTAI scores, the difference in pretest – post-test gains in scores was non-significant.

Table V shows the pretest— post-test difference scores from the Gowan Teacher Prognosis Scale of the MMPI (4) for both groups and the t values of the differences in means. The post-test scores were lower for the control group and higher for the experimental group, and the difference in means was significant at the .01 level of confidence. This finding suggests a better prognosis for teaching success for the television group. When the scores for this group were considered by ability levels, the more favorable prognosis was confined to the low seven (Wilcoxon=1. $P < .05$).

Table VI shows the pretest — post-test difference scores from the Welsh A scale of the MMPI (11) for both groups and the t values of the difference in means. All differences were non-significant. It seems that there were no noticeable changes in anxiety, as measured by this scale, for either group at the end of the semester.

TABLE I

EXPERIMENTAL AND CONTROL GROUP ACHIEVEMENT SCORE COMPARISONS FOR ALL UNITS COMBINED

N = 14 PAIRS

	Lecture Test		Test on Observations		Observation Points	
Group	Mean	t	Mean	t	Mean	t
Con	470.57		184.29		639.00	
		.39		2.07		2.49*
Exp	466.29		173.07		702.00	

* Significant at the .05 level

TABLE II

EXPERIMENTAL AND CONTROL GROUP FINAL TEST SCORE COMPARISONS

N = 14 PAIRS

	Final Lecture Test		Final Test on Observations	
Group	Mean	t	Mean	t
Con	132.43		55.93	
		.12		1.21
Exp	132.71		51.93	

TABLE III

COMPARISONS OF EXPERIMENTAL AND CONTROL GROUP SCORES ON THE FINAL EXAMINATION GIVEN FOUR MONTHS AFTER THE STUDY SEMESTER

N = 14 PAIRS

	Final Lecture Test		Final Test on Observations	
	Mean	t	Mean	t
Con	125.08		51.75	
		.08		.36
Exp	125.46		50.17	

TABLE IV

COMPARISON OF EXPERIMENTAL AND CONTROL GROUP GAINS IN MTAI SCORES
N = 14 PAIRS

Group	M	Diff.	t	S^2	t	r
Con	14.00	6.21	.92	353.83	.52	.23
Exp	20.21			473.06		

TABLE V

COMPARISON OF EXPERIMENTAL AND CONTROL GROUP PRETEST—POST-TEST DIFFERENCE SCORES OF THE TEACHER PROGNOSIS SCALE OF THE MMPI
N = 14 PAIRS

Group	M	Diff.	t	S^2	t	r
Con	-2.00	6.21	3.18*	33.52	.63	.06
Exp	4.21			23.33		

* Significant at the .01 level

TABLE VI

COMPARISON OF EXPERIMENTAL AND CONTROL GROUP PRETEST—POST-TEST DIFFERENCE SCORES OF THE WELSH A SCALE OF THE MMPI
N = 14 PAIRS

Group	M	Diff.	t	S^2	t	r
Con	.14	1.72	1.04	23.23	.95	-.04
Exp	-1.58			13.47		

Conclusions

An analysis of achievement test scores warranted the following conclusions:

1. Televised lectures were as effective as conventional instruction in bringing about increases in knowledge as measured by teacher-made objective tests.

2. For intact groups, televised observation of classroom procedures was as effective as direct observation as measured by teacher-made objective tests, but less effective than direct observation when students of high ability were considered separately.

3. Using observation points from daily papers as a criterion, televised observation of classroom procedures was more effective than direct observation for both intact groups and students of low ability, when ability groups were considered separately.

4. Televised instruction and observation of classroom procedures were as effective as conventional instruction and direct observation of classroom procedures when final grades were used as a criterion.

5. Televised instruction was as effective as conventional instruction when the final examination, given four months after the end of the semester, was used as a criterion of retention.

An analysis of pretest— post-test difference scores from the Minnesota Teacher Attitude Inventory, the Teacher Prognosis Scale, and the Welsh A Scale of the Minnesota Multiphasic Personality Inventory seemed to justify the following conclusions:

1. Instruction by television was as effective as conventional instruction in bringing about changes in teacher attitude, as measured by MTAI scores.

2. Students in the television group improved in teacher prognosis, as measured by the selected scale of the MMPI, while those in the conventional class did not. This was found for intact groups but limited to the low ability group when ability groups were considered separately.

3. Students in the television group did not show signs of anxiety, as measured by the selected scale of the MMPI, that might be attributed to receiving all instruction via television.

It is feasible to use low-cost industrial-type vidicon closed-circuit television under the conditions of this experiment.

REFERENCES

1. Carpenter, C.R., and Greenhill, L.P. "An Investigation of Closed-Circuit Television for Teaching University Courses," Instructional Television Research, Project I (University Park, Pennsylvania: The Pennsylvania State University, July 1955).
2. Carpenter, C.R., and Greenhill, L.P. "An Investigation of Closed-Circuit Television for Teaching University Courses," Instructional Television Research, Project II (University Park, Pennsylvania: The Pennsylvania State University, Spring 1958).
3. Erickson, Clifford G., and Chausow, Hymen M. The Chicago City Junior College Experiment in Offering College Courses for Credit via Open Circuit Television (Chicago: Chicago City Junior College, 1959).
4. Gowan, J.C., and Gowan, May S. "A Teacher Prognosis Scale for the MMPI," Journal of Educational Research, XLIX (September 1955), p. 4.
5. Kanner, Joseph H., et al. Television in Army Training: Evaluation of Television in Army Basic Training (Washington, D.C.:Department of the Army, Office of Chief Signal Officer, August 1958), reprint from Hum RRO Technical Report 14.
6. Kumata, Hideya. Teaching Advertising by Television, Study II (East Lansing, Michigan: Michigan State University, Communications Research Center, Winter 1957-59), mimeographed.
7. Macomber, F.G., and Siegel, Laurence. Experimental Study in Instructional Procedures, A Study of Load Credit for TV Teaching (Oxford, Ohio: Miami University, October 1958).
8. Manley, John R. Summary of TV Research at Lowry Air Force Base (Denver, Colorado: Lowry Technical Training Center, mimeographed.
9. Minnesotan, The. "Studies, Tests Reveal Teaching Potential of Closed Circuit TV," (Minneapolis, Minnesota: University of Minnesota, November 1959).
10. Rock, Robert T., Jr., Duva, James S., and Murray, John E. Instructional Television Research Reports, Technical Report No. NAVTRADEVCEM 20-TV-4. NAVEIGS P-1544 (Port Washington, Long Island, New York: U.S. Naval Training Device Center, 1956).
11. Welsh, George Schlager, and Dahlstrom, W. Grant. Basic Readings on the MMPI in Psychology and Medicine (Minneapolis, Minnesota: University of Minnesota Press, 1956).

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 3, November 1962)

Bilingualism in Southwest Louisiana

HAROLD D. LOVE
Nicholls State College
Thibodaux, Louisiana

THERE HAS BEEN considerable discussion and controversy among educators regarding the effects of bilingualism on the mental, emotional, social, and educational growth of the individual. It is widely assumed that children who learn two languages simultaneously during the process of language development are inferior to monolingual children in many areas of linguistic achievement.[1]* The problem is of special importance to those educators dealing with the large French-English bilingual population in the southwestern section of Louisiana. The present study was undertaken in order to examine the relationship between bilingualism and auditory discrimination skills.

Statement of the Problem

It has been noted that a great many children in Southwest Louisiana omit the th sound in their conversational English and also in their oral reading. This study sought to determine if the children omit the th sound because they cannot hear it due to their language background, to determine whether they can distinguish between related similar and dissimilar sounds, and to determine what relationship there is between the examiner's language background and the results of the test.

Population Studied and Procedures Used

The groups studied totaled 100 children in elementary schools in the vicinity of Southwest Louisiana. They were chosen from ten public and four private schools of this area. The children were classified into four types. The first classification included children who spoke French and English and who were tested by an examiner who spoke French and English. The second was made up of children who spoke French and English and were tested by an examiner who spoke only English. The third group was made up of children who spoke only English and were tested by an examiner who spoke French and English. The fourth was based upon children who spoke only English and were examined by a tester who spoke only English. Twenty-five children were tested in each category. An experimental form of the Wepman Auditory Discrimination Test[2] was given to every child and a specially devised th-t Auditory Discrimination Test was administered to every child.

The th-t test was made up of forty paired words. One word was pronounced correctly and the other word was pronounced without the th sound. An example of this is: thimble, timble; those, dose; mother, mudder; thumb, tumb; another, anudder. These words were alternated; the correct word being first one time and second the next time. There was an even distribution of voiced and voiceless th words.

A group of examiners who spoke French and English administered an experimental form of the Wepman Auditory Discrimination Test and the th-t test to 25 children who spoke French and English. A group of examiners who spoke only English then gave the same battery of tests to 25 children who spoke only English. French-speaking examiners gave the test to 25 English-speaking children and English-speaking examiners administered the tests to 25 French-speaking children. This is presented in Table I. All of the examiners administered the tests in English.

Findings

Table II presents a statistical analysis of the findings. It can be seen that the subjects did better on both tests regardless of what they spoke at home if the examiner spoke only English. This was significant at the .05 level. The children scored higher on the Wepman test than on the th-t

---

* Footnotes will be found at the end of the article.

TABLE I

TABLE OF TESTS AND EXAMINERS

			A			
			Tests of Auditory Discrimination			
B			Wepman Test A1	Mean Score	Th-T Test A2	Mean Score
B_1	Children who speak French and English at home	C_1	Scores on test administered by French-speaking examiner	34.5	Scores on test administered by French-speaking examiner	28.0
		C_2	Scores on test administered by English-speaking examiner	35.4	Scores on test administered by English-speaking examiner	36.1
B_2	Children who speak English at home	C_1	Scores on test administered by French-speaking examiner	31.2	Scores on test administered by French-speaking examiner	33.3
		C_2	Scores on test administered by English-speaking examiner	35.6	Scores on test administered by English-speaking examiner	34.2

TABLE II

STATISTICAL ANALYSIS OF THE FINDINGS

	SS	DF	MS
Total	3483	199	
Between	1722	99	
B	.61	1	.61
C	490	1	490*
Error	1221	96	12.71
Within	1761	100	
A	80	1	80**
BA	144	1	144**
CA	552	1	552**
ABC	443	1	443**
Error	542	96	9.56
Total	3483	199	

* Significant at the 5% level
** Significant at the 1% level

TABLE III

TEST, EXAMINER, AND SUBJECT COMBINATIONS IN RANK ORDER ACCORDING TO SCORES MADE BY ONE HUNDRED SUBJECTS

Test	Examiner's Language Background	Subject's Language Background	Rank
th-t	English	French and English	1
Wepman	English	English	2
Wepman	English	French and English	3
Wepman	French and English	French and English	4
th-t	English	English	5
th-t	French and English	English	6
Wepman	French and English	English	7
th-t	French and English	French and English	8

test regardless of the examiners. This was significant at the .01 level. The interaction between the auditory test scores and the language spoken was also significant at the .01 level. The test results depended, therefore, on the number of languages spoken at home. The effect of the language the examiner used was also significant at the .01 level. None of the variables operate independently of each other. The effect of one is influenced by the other two.

There were 100 Wepman tests and 100 th-t tests, divided into eight sets of tests. By listing the tests from one through eight, one can see how the students ranked from the lowest scores to the highest scores. This is presented in Table III.

Summary and Conclusions

Many children in southwestern Louisiana omit the th sound in their conversational English and also in their oral reading. This study sought to determine whether or not the children could discriminate between the th sound and the t sound, and whether they could distinguish between other related similar and dissimilar sounds.

Fifty children were given an experimental form of the Wepman test of Auditory Discrimination and a th-t test of auditory discrimination by examiners who had a language background of French and English. Another fifty children were administered the same tests by examiners whose language background was only English. This study sought to investigate the relationship between the examiner's language background and the results of the tests and to determine which group of children can best discriminate between these sounds in words.

It appears from this study that Southwest Louisiana children are more adept at discriminating between words when the examiner speaks only English. It also indicates that the children in this area who speak French and English and the children who speak only English tend to do better on the Wepman test of many sounds than on the th-t test of only one sound. The subjects who speak French and English score higher on the Wepman test than do the children who speak only English, which suggests that their general auditory discrimination skills are adequate. But the one-language group does better on the th-t test than the two-language group. Success on these tests is apparently influenced by the type of examiner and the number of languages spoken at home. They do not operate independently of each other and the effect of one is influenced by the other two.

The children from a French-English language region apparently do hear the th sound and they can discriminate between it and the t sound better when spoken to by an examiner with only an English language background.

FOOTNOTES

1. Carrow, M. A. "Linguistic Functioning of Bi-Lingual and Mono-Lingual Children," Journal of Speech and Hearing, pp. 371-80.
2. Some of the paired words from both forms of the Wepman Auditory Discrimination Test made up this experimental form. The Wepman test can be bought from Joseph M. Wepman, 950 East 59th Street, Chicago 37, Illinois.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 3, November 1962)

A Study of the Relationship Between Parental Acceptance and the Academic Achievement of Adolescents

JANICE M. BARWICK
Hofstra College

DUGALD S. ARBUCKLE
Boston University

Many studies have been carried on showing the relationship between parental attitudes and variations in the personality of the child. Research has demonstrated that negative parental attitudes, whether representing overt hostility, ambivalence, or a neutral, passive relationship between child and parents, may result in the following behavioral manifestations in children: 1) feelings of insecurity and extreme sensitivity to attention (5, 6); 2) shy, withdrawn, submissive behavior (2); or 3) aggressive or dominant behavior (3, 8).

Relatively little research has been carried on which concerns itself with the influence of parent attitudes on academic achievement (1, 7), although in reality achievement is one manifestation of personality. A student who fails to achieve up to capacity may be a student who is not motivated to do so, and one of the most important factors influencing the motivation of children is parental attitudes. Abraham Maslow of the Nebraska Symposium in 1955, explained this motivation towards achievement which results from parental acceptance in terms of self-actualization (4): "So far as motivational status is concerned, healthy individuals have sufficiently gratified their basic needs for safety, belongingness, love, respect, and self-esteem so that they are motivated primarily by trends toward self-actualization."

The present study, then, was undertaken to determine whether there is a significant relationship between parental acceptance,as perceived by both parents and children, and the academic achievement of the latter.

Procedure

The sample population was selected from students enrolled in Grade 8 in a junior high school in a suburban community of Massachusetts. The group was classified into high, average, and low achievers on the basis of grade-point indices. High and low achievers were identified on the basis of grade-point indices one standard error of estimate above or below the predicted value. The predicted grade-point index was found by means of a regression equation, using an intelligence quotient obtained from the Otis Beta Group Intelligence Test as the prediction variable. The grades received at the termination of the first semester of Grade 8 were used to compute the actual grade-point index. This procedure resulted in 47 high achievers, 226 average achievers, and 37 low achievers. From each of these groups 15 boys and 15 girls were selected according to the method of random numbers, resulting in 30 students at each level. Those students and their parents comprised the sample population.

The instruments used to measure the children's perception of parental acceptance were: Whitesel's Situation Questionnaire for Parent-Child Relationships, Gilmore's Sentence Completion Test, and the TAT. The Porter Parental Acceptance Scale was used to obtain data concerning the parents' perception of their acceptance of their children. Each of the tests yielded three scores: Mother-Acceptance, Father-Acceptance, and Total Parent-Acceptance Scores. The data were subjected to an analysis of variance and the results of each test were treated separately. The level of significance chosen was the five percent level.

Results

The results reported are based on the trends indicated by mean values on the assumption that trends in mean values of the mother- and father-acceptance scores for the three groups of achievers would indicate that a relationship exists between parental acceptance and academic achievement. The names of the tests which yielded mean values demonstrating these trends are reported, and wherever trends are supported by significant F-ratios, the names of the tests and the level of significance are indicated.

1. Boys achieving at a higher level reported fathers as more accepting than did those achieving at a lower level. (Table I, Situation Questionnaire and the Sentence Completion Test).
2. Both the average and low achieving boys perceived their mothers as more accepting than did the high achieving boys. (Table I, Situation Questionnaire, Sentence

TABLE I

TABLE OF MEANS OF MOTHER- AND FATHER-ACCEPTANCE SCORES OF THE HIGH-, AVERAGE-, AND LOW-ACHIEVING BOYS AND OF THE MOTHERS AND FATHERS OF THE HIGH-, AVERAGE-, AND LOW-ACHIEVING BOYS

(N=15)

Instruments	Mother-Acceptance			Father-Acceptance		
	High	Average	Low	High	Average	Low
Situation Questionnaire	6.6	14.1	13.2	6.3	3.2	-.83
Sentence Completion Test	.20	.60	.13	.66	.53	.06
Thematic Apperception Test	.37	-.64	-.58	-.43	-1.14	-.33
Parental Acceptance Test	137.3	137.5	139.8	128.7	141.3	134.7

TABLE II

TABLE OF MEANS OF THE MOTHER-VERSUS THE FATHER-ACCEPTANCE SCORES OF THE HIGH-, AVERAGE-, AND LOW-ACHIEVING BOYS AND OF THE MOTHERS AND FATHERS OF THE THREE GROUPS

(N=15)

Instruments	Boys					
	High-Achieving		Average-Achieving		Low-Achieving	
	Mother	Father	Mother	Father	Mother	Father
Situation Questionnaire	6.6	6.3	14.1	3.2	13.2	-.83
Sentence Completion Test	.20	.66	.60	.53	.13	-.06
Thematic Apperception Test	-.37	-.43	-.64	-1.14	-.58	-.33
Parental Acceptance Scale	137.3	128.7	137.5	141.3	139.8	134.7

TABLE III

TABLE OF MEANS OF MOTHER-AND FATHER-ACCEPTANCE SCORES OF THE HIGH-, AVERAGE-, AND LOW-ACHIEVING GIRLS AND OF THE MOTHERS AND FATHERS OF THE HIGH-, AVERAGE-, AND LOW-ACHIEVING GIRLS

(N=15)

Instruments	Mother-Acceptance			Father-Acceptance		
	High	Average	Low	High	Average	Low
Situation Questionnaire	15.2	13.4	11.5	5.5	3.3	.27
Sentence Completion Test	.60	.26	.33	.80	.46	.33
Thematic Apperception Test	-.18	-.38	-.58	-.20	-.20	-.73
Parental Acceptance Scale	141.3	141.6	130.9	133.1	130.7	140.1

TABLE IV

TABLE OF MEANS OF THE MOTHER-VERSUS THE FATHER-ACCEPTANCE SCORES OF THE HIGH-, AVERAGE-, AND LOW-ACHIEVING GIRLS AND OF THE MOTHERS AND FATHERS OF THE THREE GROUPS

(N=15)

Instruments	Girls High-Achieving Mother	 Father	 Average-Achieving Mother	 Father	 Low-Achieving Mother	 Father
Situation Questionnaire	15.2	5.5	13.4	3.3	11.5	.27
Sentence Completion Test	.60	.80	.26	.46	.33	.33
Thematic Apperception Test	-.18	-.20	-.38	-.20	-.58	-.73
Parental Acceptance Scale	141.3	133.1	141.6	130.7	130.9	140.1

TABLE V

TABLE OF F-RATIOS OBTAINED FROM THE ANALYSIS OF VARIANCE OF THE MOTHER- AND FATHER-ACCEPTANCE SCORES

(N=45)

Instruments	Mother-Acceptance Scores Boys	 Girls	Father-Acceptance Scores Boys	 Girls
Situation Questionnaire	2.52	.44	1.32	.74
Sentence Completion Test	.95	.24	1.55	.65
Thematic Apperception Test	.38	1.5	2.93	4.60*
Parental Acceptance Scale	.12	2.99	1.03	1.8

* Indicates the five percent level of significance

TABLE VI

TABLE OF F-RATIOS OBTAINED FROM THE ANALYSES OF VARIANCE OF MOTHER- VERSUS FATHER-ACCEPTANCE SCORES

(N=15)

Groups	Situation Questionnaire	Sentence Completion Test	Thematic Apperception Test	Parental Acceptance Scale
High-Achieving Boys	.004	.05	.06	.53
Average-Achieving Boys	4.31	.04	4.91*	.73
Low-Achieving Boys	10.25**	.12	.47	.90
High-Achieving Girls	5.73*	.29	0.00	1.90
Average-Achieving Girls	5.31	.81	.78	4.40
Low-Achieving Girls	2.87	0.00	.47	4.70*

* Indicates the five percent level of significance
** Indicates the one percent level of significance

Completion Test). This increase in mother acceptance with a decrease in father acceptance as the level of achievement became lower may be inferred as an attempt by the mothers of the lower achieving boys to compensate for the decrease in father acceptance.

3. a. High-achieving boys perceived their mothers and fathers as equally accepting (Table II, Situation Questionnaire, Sentence Completion Test, TAT).
   b. Average-achieving boys perceived their mothers as more accepting than they did their fathers (Tables II and VI, Situation Questionnaire, TAT - five percent level of significance).
   c. Low-achieving boys perceived their mothers as more accepting than they did their fathers (Tables II and VI, Situation Questionnaire - one percent level of significance, Sentence Completion Test.
4. Father acceptance as perceived by the girls increased as the level of achievement became higher (Tables III and V, TAT - five percent level of significance, Situation Questionnaire, Sentence Completion Test).
5. Mother acceptance as perceived by the girls increased as the level of achievement became higher (Table III, Situation Questionnaire, TAT).
6. Mother acceptance was perceived by the girls at all levels as greater than father acceptance, on the basis of the results of the Situation Questionnaire, which were significant at the five percent level for both the high and average-achieving girls (Tables IV and VI). The data of the other tests produced no significant difference between mother and father acceptance as perceived by the girls.
7. In general, no particular trends were evident in the mean values of the scores of tests measuring the parents' report of their own attitudes. However, it is felt that the child's perception is a stronger factor in achievement than the actual stimulus or the parental report of it.

The results of this study would indicate that, in attempting to solve the problem of underachievement in our schools, it would be advisable for the teacher of the underachiever to discuss the problem with the student and his parents in order that all concerned may understand the dynamics of the problem. Furthermore, it is obvious that effective counseling services for both children and parents would assist greatly in preventing and solving the problem. All too often, the schools and the parents rely upon "warning" or "deficiency" notices, threats of withdrawal from athletic teams, band, and other extracurricular activities, in the belief that such steps will force the student to achieve to capacity. Many times, these very steps aimed at ameliorating the situation force the child deeper into a state of underachievement. It should be pointed out that a brief, perfunctory conversation between counselor and student, which in many schools is a ritual which follows the issuance of "warning" notices, may ease the conscience of the counselor but will in itself lend little to the solution of underachievement. In conclusion, one must realize that the frequency of underachievement in our schools is great, and that the probability of successfully solving the problem at the secondary level is much less than during the first few years of school. Consequently, earlier identification of the underachieving student and the availability of counseling services at the lower grade levels for parents and students are most important.

REFERENCES

1. Gilmore, John V. "A New Venture in the Testing of Motivation," The College Board Review, XV (1951), pp. 221-26.
2. Gottemoller, R. "The Influence of Certain Aspects of the Home Environment on the Adjustment of Children to Kindergarten," Smith College Studies in Social Work, IX (1939), pp. 303-59.
3. Lewis, Hilda. "Unsatisfactory Parents and Psychological Disorders in Their Children," Eugenic Review, XLVII (1955), pp. 153-62.
4. Maslow, Abraham. "Deficiency Motivation and Growth Motivation," Nebraska Symposium on Motivation, Marshall R. Jones, ed. (University of Nebraska Press, 1955), p. 8.
5. Newell, H. W. "Psycho-dynamics of Maternal Rejection," American Journal of Orthopsychiatry, IV (1934), pp. 387-403.
6. Symonds, Percival. Psychology of Parent-Child Relationships New York: Appleton-Century Company, (1939).
7. Tibbetts, John R. "The Role of Parent-Child Relationships in the Achievement of High-School Pupils," Ph.D. Dissertation, New York University (1954).
8. Zimmerman, A. "Parental Adjustment and Attitudes in Relation to the Problem of Five- and Six-Year-Old-Children," Smith College Studies in Social Work, I (1931), pp. 406-7.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 3, November 1962)

Teachers' Ratings of High-School Students on "Likability" and Their Relation to Measures of Ability and Achievement

J. R. WILLIAMS and WALTER W. KNECHT
Public Schools, Kankakee, Illinois

ONE OUTGROWTH of the quickened interest in education during the last five years has been the active serarch for better means of predicting school achievement. In this quest, personality variables are receiving extensive consideration by several investigators (2, 4, 6, 7, 11, 12). Generally, these efforts have been rewarding either through discovering a significant correlation between certain personality measures and level of achievement or by finding that some combination of personality and ability measures gives a higher correlation with the criterion than does ability test scores alone.

Judging from the nature of the relationships so far disclosed between personality traits and achievement, one might expect that accurate prediction of achievement is not possible unless a fairly large number of personality test or "factor" (in addition to ability) scores are available for each student. If this is true, then the accumulation and analysis of test data necessary for prediction could conceivably be a forbidding task to many school systems. Under such circumstances, a more concise means of personality assessment, if equally reliable and valid, would be welcomed. The study reported here represents an attempt by the authors to investigate some aspects of the scientific worth of one such method.

In spite of the tang of antiquity and the often disappointing results associated with rating techniques, this method was chosen for study. In a previous investigation (13), the senior author found psychiatric ratings of young adults on a personality variable to be fairly reliable under conditions where there was considerable breakdown and explanation of the rating task. The rating of students by their teachers did not seem to present a problem of greater scientific magnitude. In fact, some advantage seemed to exist from the fact of long and close acquaintance usually found in the teacher-student situation.

To continue with this "telescopic" approach to personality measurement, it seemed necessary to choose a rather broad and familiar trait for rating. In addition to being a "recognizable composite," it was equally desirable, for the ulterior purposes of this study, that the trait have something to recommend it as a hypothetical influence on school achievement. While pondering these considerations, the trait of "likability" came to mind.

The reality and positive influence of a trait similar to that of "likability" is not hard to document in the experience of most people. A pleasing manner seems to contribute to one's effectiveness whether he be a salesman, minister, teacher, or politician. And, how many remember some fellow student from school days who seemed to please teachers to the point of getting a little extra consideration, privilege-wise as well as grade-wise?

The students who seemed to please teachers and make good grades in high school seemed also to continue to be successful in college. Were these students stupid? No—stupidity is seldom pleasing to anyone. Were they intelligent? Yes—at least average and, perhaps, above average in most cases. Are all intelligent students pleasing or "likable" to teachers? Certainly not! Maybe intelligence (ability) and "likability" are both factors in school achievement, each supplementing the other.

The present investigation, conceived as preliminary to the larger task of methodological validation or prediction, was designed primarily to help answer the following questions:

1. To what extent are teachers' ratings of students on the trait of "likability" in agreement?
2. To what extent are these same ratings related to student ability and achievement?

Procedure

The plan of the study called for getting a representative sample of students of known ability and achievement and obtaining teacher ratings of these students on the trait of "likability." Then, relationships between these ratings and the students' ability and achievement scores had to be determined.

The senior class of high school was selected as the sample of students to be rated, the main reason being that the seniors, as a whole, were known by a relatively large number of teachers and had been in school long enough for sufficient measures of ability and achievement to have been obtained on them.

The rating procedure was intoduced to the teachers by means of a one-page letter. In this communication it was stated that there was interest in "improving our ability to predict academic achievement of students." Mention was made of the recognized importance of ability as a factor in achievement and the teachers were told of the belief that knowledge of personality factors would increase one's ability to predict academic grades. The letter continued as follows:

"At this point we are not satisfied that we are measuring all the important variables needed in order to get good prediction in many individual cases. Therefore, we would like to have you help us obtain a rating of what we feel is an important personality variable.

"The variable we would like to have you rate is one that we are calling 'likability' or 'likeableness.' For this study, we will define likability as the combination of personality traits or factors (physical, social, emotional) which tends to give others a favorable impression of the person.

"We are asking you to rate each senior student you have had in class on a scale of likability. To aid your understanding of what we want you to do we are giving an example:

Name John Doe

Degree of Likability

Very Little	Below Average	Average	Above Average	Very Much
		X		

"We have rated John Doe on likability by making a check (X) along the scale to indicate that we think he is nearly average in degree of likability. Note that you can place the check anywhere along the line and not just at the scale divisions. The descriptive terms and division points are merely guides to indicate relative degrees of likability.

"Do not sign the rating sheets and take care that no student sees them. Please be as accurate as you can in your ratings......"

Beginning as in the example above, the rating sheets contained the names of all the seniors (N = 371) and, opposite each name, a line with five equally-spaced division points. All lines were of the same length, 110 millimeters, and occupied relatively the same positions to the right of the names. The numerical value of any rating was determined as the measured distance in millimeters from the left, or zero, end of the respective line to the check (X) appearing along it. The average of these values for any subject became his raw score or degree of "likability."

Extent of rater agreement was determined by the split-half method, making use of the Spearman-Brown correction formula. In all cases, where there were four or more ratings, these ratings were divided into two halves (first and second) and the average of each half calculated. If a subject had an odd number of ratings, the middle one was omitted in the division process.

Total scores on the California Test of Mental Maturity (CTMM) and "B Factor" scores on the 16 P. F. Test (3) were available as ability measures. The grade point averages at the end of three and a half years of high school were used as achievement scores. All scores ("likability," ability, and achievement) were standardized previous to computing the Pearson product-moment correlations among them.

Results

In this section, the results will be given in the order of their relation to the main purposes of our investigation. First, with regard to the reliability of the teacher ratings, the value of r was .557 and, after correction, .716. Considering the fact that in the case of 101 subjects there were only six or less ratings, this value is deemed an indication of respectable reliability.

TABLE I

CORRELATIONS AMONG LIKABILITY RATINGS, ABILITY SCORES AND GRADE POINT AVERAGES

	Likability Ratings	CTMM Scores	16 P. F. "B Factor" Scores
CTMM Scores	.529		
16 P. F. "B Factor" Scores	.478	.752	
Grade Point Averages	.668	.565	.550

Next, Table I shows the intercorrelations among "likability" ratings, ability scores, and grade point averages. The main results can be enumerated as follows:

1. Positive relationships, significant at less than the .01 level of confidence (5), exist among "likability" ratings, ability scores, and grade point averages of the students.

2. The highest correlation is that between the two measures of ability.

3. "Likability" ratings correlate higher with

grade point averages than do ability scores.

The last result is the one of major concern here. It signifies that, under the conditions of this study, a student's achievement in terms of academic grades is more closely associated with his degree of "likability," as rated by teachers, than with his ability, as measured by tests.

Discussion

This study was prompted by what seemed to be a need in the schools for a convenient, yet scientifically appropriate, method for measuring personality variables believed to be related to student achievement. The fact that this preliminary investigation has produced some positive and encouraging results is worthy of further comment.

First, what can one conclude from this study about the reliability of teacher ratings? One answer would be that the present results ($r = .71$) agree favorably with those reported by others under somewhat comparable conditions. For instance, a more comprehensive study in terms of number of variables and classes of students but involving smaller populations (10) gave a range of reliabilities for teacher ratings from .58 to .83. Differences in reliabilities seemed to be mostly due to the nature of the traits being rated.

A further analysis of the ratings obtained in the present investigation shows clearly that reliability is a function of the number of ratings. For that group of students where there were six or fewer ratings the reliability was only .63, compared with .71 for the total group. A safe conclusion from this limited study seems to be that teacher ratings are reasonably dependable, particularly when six or more ratings are available for split-half division and correlation.

Next, what meanings can be attributed to the observed correlations among "likability" ratings, ability scores, and grade point averages? In the first place, the results here seem at least roughly consistent and logical in the sense that both ability measures correlate about equally well with "likability" ratings and with grade point averages, respectively, and that the highest correlation is that between the two ability measures themselves. These facts give some assurance of the validity of the ability measures used.

With respect to the relations of ability scores and school achievement, the results of this study are in close agreement with those of some earlier investigators (1, 8). The statement of this relationship is usually to the effect that ability scores account for approximately one-fourth of the achievement variance.

Perhaps the finding in this study that most challenges interpretation is the rather high correlation between "likability" ratings and grae point averages. It is mildly surprising, to say the least, to find that a supposedly "nonintellectual" measure correlates higher with school grades than does an "intellectual" one. A possible clarification exists in the fact that ability scores have approximately the same relationship to both "likability" ratings and grade point averages. This suggests that a common factor may underlie the three of them. But the nature of any such factor is not at once available from the data. Ryan, in the article referred to above, seemed puzzled over the same or similar observations. He states:

> It may be that ratings represent very little more than imperfect reflections of academic performance—i. e., they may be given 'after the fact' of grades. Yet, at least to some extent, trait ratings may tap personality variables that do affect academic achievement. (10:128)

The concluding and reserved statement of Ryan's conforms to the speculation offered by the writers in the introduction. This is to the effect that both ability scores and "likability" ratings may be able to make separate contributions to the task of predicting school achievement. With this in mind and using the relevant relationships in Table I, it was possible to obtain the following standard score form of the multiple regression equation (9:147):

$$z_1' = .307z_2 + .540z_3 \quad ,$$

where z_1' represents predicted grade point average, z_2 stands for ability score, and z_3 is the "likability" rating for any student. This equation, of course, is only a first approximation to accurate prediction, since there is need for improving the measure of "likability" and, more important, there are other, undoubtedly important variables that will have to be included. A repetition of this study, using a different sample of students, would offer an opportunity to remove some of these shortcomings and to test the predictive power of the above equation.

REFERENCES

1. Cattell, R. B. Measured Personality Factors in School Children and Their Relation to Achievement and Creativity, abstract of address before Spring Meeting of Illinois Psychological Association, 1959.
2. Cattell, R. B. Predicting School Achievement from Personality as Measured by the 'HSPQ' and '16PF' Tests, Information Bulletin, No. 4 (Champaign, Ill.: Institute for Personality and Ability Testing, 1960).
3. Cattell, R. B., Saunders, D. R., and Stice, G. Handbook for the Sixteen Personality Factor Questionnaire (Champaign, Ill.: Institute for Personality and

4. Cattell, R. B., Beloff, H., and Coan, R. W., Handbook for the IPAT High School Personality Questionnaire (Champaign, Ill.: Institute for Personality and Ability Testing, 1958).
5. Edwards, A. L., Statistical Analysis (New York: Rinehart and Co., 1946).
6. Holland, J. L. The Prediction of College Grades and Post College Achievement, Technical Report No. 4, Research and Studies (Evanston, Ill.: National Merit Scholarship Corporation, 1960).
7. Holland, J. L. "The Prediction of College Grades from Personality and Aptitude Variables," Journal of Educational Psychology, LI (1961), pp. 245-54.
8. Holzinger, K. J., and Crowder, N. A. Manual, Holzinger-Crowder Unifactor Tests (Yonkers-on-Hudson, N. Y.: World Book Co., 1955).
9. McNemar, Q. Psychological Statistics (New York: John Wiley and Sons, 1949).
10. Ryan, F. J. "Trait Ratings of High School Students by Teachers," Journal of Educational Psychology, XLIX (1958), pp. 124-28.
11. Quinn, S. B. "Relationship of Certain Personality Characteristics to College Achievement," Dissertation Abstracts, XVII (1957), 809 pp.
12. Warren, J. R., and Heist, P. A. "Personality Attributes of Gifted College Students," Science, CXXXII (1960), pp. 330-37.
13. Williams, J. R. "A Test of the Validity of P-Technique in the Measurement of Internal Conflict," Journal of Personality, XXVII (1959), pp. 418-37.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 3, November 1962)

Grade Placement of High-School Biology

H. SEYMOUR FOWLER

Pennsylvania State University

MUCH HAS been said about the appropriate grade placement of biology. We also find much written. Some of the better articles are indicated in the references. Sharkan (5) indicates, in general terms, the pros and cons of presenting biology at the ninth-grade level. His conclusions bring out the need for additional research to substantiate placement of biology at either the ninth-grade level or at the tenth-grade level.

One of the early articles on the subject was by Goldstein (1), who concludes: "I do not think that high-school freshmen are mature enough to study biology, at least not the biology we teach in New York City." He also reports that on test scores in the New York State Regents Examination, his classes scored as follows: In his ninth-year biology course, ten out of 32 of the pupils (31.3 percent) scored above the median of the total group, while in his tenth-year biology course, 25 out of 38 scored above the median of the total group.

Matala (3) gives opinions on when to teach biology, and also suggests questions which must be answered: "I doubt that anyone will question that a biology course can be taught ninth graders with success, or that ninth graders can do as well as tenth graders in many cases and with many types of courses. But there are questions that I would like to raise--loudly and long. Is this the biology we want the potential scientists to have?"

Mathes and Blanc (4) reported on a study which involved most of the students taking biology in junior and senior high school during the second semester of the 1958-59 school year in Denver, Colorado schools. Their findings: "point out clearly that the group of ninth graders selected on the basis of ability, interest in science, and teacher's recommendations did considerably better than unselected high school students. When compared with a group of tenth-grade students selected to match the ability of the ninth grade, ninth-grade students did slightly better than the tenth-grade group."

Heidegerd (2) reported on his doctoral thesis completed at the University of Illinois. His conclusions were: "It would seem that the change to ninth-grade biology and tenth-grade physical science should only be made if it is expected that there will be some loss in biology achievement and that other gains will make the change worthwhile."

After reading a variety of studies and bringing together all findings, it would appear that more research is needed. Hence, the following report is given to add additional information which may prove useful in helping a school system to decide on the grade placement of high-school biology.

This study involves the administration of an achievement test to students in two different types of situations. The first group was made up of ninth graders who had completed a year of biology. The second group was made up of tenth graders who had completed a year of biology. Schools in which the tests were given are in central and eastern Pennsylvania. Because of the predominance of the 6-3-3 plan of organization of our schools, it is difficult to find teachers who are responsible for both ninth- and tenth-grade biology. This is one problem which could not be overcome in designing the research project. In other words, it was not possible to compare performance of students, both ninth graders and tenth graders, working with the same teacher.

The Cooperative Biology Test, Form Y, distributed by the Educational Testing Service, was the achievement test used in the study. Test forms, answer sheets, and instructions were mailed to each cooperating teacher in the early spring months prior to the June testing period. In June, the tests were administered by each teacher. The completed answer sheets were mailed to the author, who scored them and mailed a copy of the results to the cooperating teachers. Data thus collected were used in the study. These data are reported in Table I.

These data serve only as introductory material. If we consider both groups as total units, then the data are, at least, interesting. We note that the ninth graders scored somewhat higher on the average on the Otis Quick Scoring Intelligence Test. However, their mean score on the Biology Achievement test was two points lower than the mean score of the tenth graders. Whether these differences are significant is problematical. Seven of the eleven classes of ninth graders were composed of

TABLE I

RESULTS OF ANSWER SHEETS

	Total Number of Students	Number of Teachers	Number of Classes	Mean IQ	Mean Test Score
9th	312	8	11	115	59
10th	340	7	13	113	61

TABLE II

RHO - RANK ORDER CORRELATION

	Rating of Teacher vs. Mean Score Achievement Test	Rating of Teacher vs. Mean IQ
9th	.7091	.2681
10th	.0001	.1676

students who were selected as superior students. The students in the four other classes were unselected. All of the tenth-grade classes were unselected, and each of the participants elected the course.

An interesting study appeared as a possibility since we did not have teachers who were presenting biology in both ninth and tenth grades. A questionnaire was produced which gave each participant an opportunity to evaluate all of the teachers under whom he had had instruction. The questionnaire made provision for the student to evaluate his biology teacher. (See copy of questionnaire following references.) All of the students completed the questionnaire twice. The interval of time between the two student evaluations was two months. A rating scale from one to nine was used for each question. In other words, "very poor" equaled one, "average" equaled five, and "very good" equaled nine.

Using these as reference points, each paper was evaluated. All scores for question one were determined, and a total was derived. The total score was divided by the number of students in the class. This score was recorded. The same procedure was repeated. To derive final scores, the average between the two determinations was taken. The reader will note that question one asks the student to evaluate all of his teachers, while two, three, and four refer to his teacher of biology. To determine the biology teacher's rating, we took an average of two, three, and four, and then subtracted the value of one from it. This gave an idea of how the student compared his biology teacher to other teachers. Let us call this a Rating-of-Teacher Score. Using these scores, we tried to determine a Rho score between this score and Mean IQ and Mean Achievement Test Scores. Table II shows the Rho scores which were determined.

These Rho scores indicate that at the ninth grade level for these classes, there seemed to be a correlation between the rating of the teacher by the students and the students' scores on the achievement test.

Rho scores for the relationship between mean IQ and Mean Score on the Achievement Test were also determined. The data were ninth grade, Rho .7540; tenth grade, Rho .1932. Again, there seems to be a correlation between the two items compared for the ninth-grade group, but very little for the tenth-grade group.

Now, refer to the mean scores on the achievement test. Let us compare the mean achievement test scores for the ninth-grade classes with the mean achievement scores for the tenth-grade classes. In this case, we employed Fisher's Method for determining the significance of the difference between two means when the results cannot be paired. In determining "t" in this case, we find there is no significant difference between the two means.

All of this information seems to substantiate the contention that ninth graders can achieve as well as tenth graders in high-school biology. One must be cognizant of the fact that a portion of the ninth-grade classes in this study were selected because of their special abilities. Hence, to be able to make a generalization concerning achievement, we would have to find an equal number of classes of unselected ninth and tenth graders and repeat the procedure. Another problem involves the teacher variable. To be able to be most exact, it would be necessary to pair classes of ninth and tenth graders under the same teacher.

From the results of this study, it would appear that we might place biology in either the ninth or tenth grade and find that our students would achieve equally well in either grade. Many teachers will say "Achievement of a test score is only one minor objective of a high-school course. Perhaps we should look to other things before making our decision about grade placement." This author agrees completely with that opinion.

REFERENCES

1. Goldstein, Philip. "Concerning Ninth-Year Biology," The Science Teacher, XXV (December 1958), pp. 454-57.
2. Heidgerd, Lloyd. "More on Ninth-Grade Biology," The Science Teacher, XXVII (March 1960), pp. 27-30.
3. Matala, Dorothy. "A Biology Course--When to Teach It," The American Biology Teacher, XXII (May 1960), pp. 270-71.
4. Mathes, George E. and Blanc, Sam. "Biology Achievement in Grades Nine Through Ten," The Science Teacher, XXVII (March 1960), pp. 23-26.
5. Sharkan, William. "Ninth Grade Biology — Pros and Cons," School Science and Mathematics, LIX (December 1959), pp. 718-22.

Teacher Number ________

Performance Scale for High School Biology Teachers

DO NOT PUT YOUR NAME ON THIS FORM.

Note: We need information about your teacher of Biology. You are asked to rank your teacher by placing an X in the box that most nearly describes your teacher or teachers.

1. How do you rank ALL the teachers you have had up to now in the school you are now attending as to their ability to make their subject matter understandable and meaningful to you?

☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐

Very Poor — Average — Very Good

2. How do you rank your teacher of Biology as to his (her) ability to make Biology understandable and meaningful to you?

☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐

Very Poor — Average — Very Good

3. How do you rank your teacher of Biology as to his (her) knowledge in the field of Biology?

☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐

Very Poor — Average — Very Good

4. How do you rank your teacher of Biology as to his (her) ability in the field of Science as a whole?

☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐

Very Poor — Average — Very Good

Fold your paper after you have marked the four checks. Place it in the envelope. Seal the envelope. Turn it in to your teacher.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 3, November 1962)

Personality Ratings of High School Students

DONALD A. LETON
University of California at Los Angeles

THE ASSESSMENT of personality characteristics through ratings has permeated nearly all activities in which the assessment of individuals is required. Although the validity of ratings can be seriously questioned, their use has been especially widespread in education. Research by Barron (1) and by Woodworth (14) indicates that ratings are an inefficient and perhaps inadequate method for the selection or description of certain student characteristics. Cattell and Drendahl (3) report inconsistencies in teachers' ratings. Their findings indicated that students often possess characteristics which are in direct contrast to those rated by the teacher.

There also are several favorable and optimistic views expressed in the literature concerning teacher ratings. Swab and Peters (12) discuss the validity and reliability of ratings for certain types of conduct in normal-life situations. Benson (2) found that teachers' ratings are of value in identifying maladjustment in the school setting.

The assessment of students by means of teacher ratings has a number of possible consequences. Holland (6) and others have pointed out that teacher ratings may effect students in various ways: e. g., in their planning for careers, in obtaining scholarships, in developing positive or negative self-concepts, and in their feelings of capability or inadequacy for achievement and creativity. Since teachers' ratings can play such an important role in the educational welfare of students, it seems appropriate to make current investigations of trait ratings by teachers. Ryan (10) and Holland (6) have suggested the use of empirical correlates to explore teacher ratings further.

The purpose of this investigation was to study the relationship among teachers' ratings and their correlation with personality scale scores. Although teachers are required to assess students' personal and social behavior as well as their achievement, their ability to perform these assessments adequately has not been seriously questioned in research.

Method

The subjects consisted of a sample of 150 boys and 115 girls who were enrolled in the ninth grade of a west-coast unified school district.* In the fall of 1958, these students were administered the California Psychological Inventory (CPI) (Gough, 5), in their Social Studies classes. Independent personality ratings corresponding to the 18 CPI scales were subsequently obtained from the tenth-grade English teachers in the fall of 1959. The tenth-grade teachers did not have any foreknowledge of the results of the personality inventory from the previous year.

The CPI is a recent but well-known instrument. It has been described in a number of recent publications (7, 4, 13, 11, 9, 8). The test consists of the following 18 subscales: Dominance (Do), Capacity for Status (Cs), Sociability (Sy), Social Presence (Sp), Self-acceptance (Sa), Sense of Well-Being (Wb), Responsibility (Re), Socialization (So), Self-control (Sc), Tolerance (To), Good Impression (Gi), Communality (Cm, Achievement via conformance (Ac), Achievement via independence (Ai), Intellectual efficiency (Ie), Psychological-mindedness (Py), Flexibility (Fx), and Femininity (Fe).

The Personality Rating Sheets which were completed by the tenth-grade teachers were derived from the CPI manual. Seven-point rating scales were presented with the title of each trait. The descriptive adjectives which are associated with the scale extremes in the manual were listed at the end of each rating scale. In the construction of the rating sheets, an effort was made to minimize central tendency. The average possession of a trait was not characterized in the description, whereas each of the extremes was somewhat balanced with the list of personal adjectives.

Results

The correlations among teacher ratings for the 264 students on the various traits were determined

---

* Appreciation is hereby expressed to the El Segundo Schools, El Segundo, California, for the permission and cooperative assistance in conducting this study.

TABLE I

INTERCORRELATIONS OF TEACHER RATINGS OF STUDENTS' PERSONAL CHARACTERISTICS

	Do	Sc	Sy	Sp	Sa	Wb	Re	So	Sc	To	Gi	Cm	Ac	Ai	Ie	Py	Fx	Fe	r
Dominance		90	85	83	72	78	71	60	47	74	71	6-	69	78	77	72	74	36	70
Capacity for Status			87	82	74	81	74	64	53	77	75	63	70	78	78	72	75	43	72
Sociability				86	76	80	69	59	45	74	74	63	69	73	74	72	75	43	71
Social Presence					74	77	68	58	43	67	70	57	64	69	72	70	71	34	68
Self-Acceptance						66	49	43	31	60	57	41	48	59	56	69	69	22	59
Sense of Well-Being							84	74	64	80	77	72	77	77	81	66	68	50	75
Responsibility								84	77	81	76	80	82	79	83	62	64	60	78
Socialization									82	80	73	82	80	72	77	54	58	67	76
Self-Control										73	66	79	77	69	68	48	55	64	75
Tolerance											83	79	83	82	79	67	70	56	78
Good Impression												78	78	74	75	67	70	56	77
Communality													85	72	75	58	63	66	77
Achievement via Conformance														78	82	60	66	61	78
Achievement via Independence															86	74	76	52	79
Intellectual Efficiency																72	74	55	79
Psychological-Mindedness																	76	37	78
Flexibility																		44	78
Femininity																			50

Decimals Omitted

Average r = 73

and the intercorrelations between ratings and scale scores were computed. Table I represents the correlation matrix for the teachers' ratings.

The average for the 18 mean intercorrelations is .73 with a range of .50 to .79. It is evident that the ratings are closely related despite the variety in the traits being rated. The high intercorrelations among the ratings suggest that there is a very strong halo effect making independence negligible. The highest and most representative variables in the matrix are those that rate the students on their achievement (Ac, Ai), on intellectual efficiency (Ie), and on intellectual interests (Py and Fx). These reflect teachers' biases toward, and agreement about, scholastic achievement and intellectual interest which are instrumental in school relationships. The low correlation (.50) between the femininity-masculinity and other ratings indicate that teachers perceive this trait as being somewhat independent of other personal-social traits.

The correlations between students' scores and teacher ratings are shown in Table II. The results indicate a significant relationship between ratings and scores with the exception of one scale (Fx). The teachers' inability to rate Flexibility may have stemmed from their lack of understanding about this trait as it is defined in the CPI. The fact that their ratings on this trait correlated highly with other ratings indicates that they may have a general concept of flexibility as it pertains to intellectual characteristics. They did not, however, conceive it as a distinct personality trait and were unable to assess it as the inventory did. Perhaps this lack of agreement can be attributed in part to a lack of semantic clarity. Nunnally and Husek's study (9a) points up the need for personality test factors to be described and interpreted in such a way that test-users can accurately understand them.

The higher and more significant correlations presented in Table II are in the scales which pertain to achievement, intellectual interest, socialization, and responsibility. The latter two CPI scales are perhaps measuring the social attributes which incline a student toward achievement in school. These findings suggest that those students who display achievement, intellectuality, sociability, and responsibility, and those who show a lack of these traits, can be rated more successfully by teachers than studetns who show variabilities in other traits.

Discussion

In similar studies, Holland (6), Ryan (10), and Talent (12a) found a strong association between teachers' ratings and academic achievement. They also listed such traits as leadership potential, self-control, and responsibility as high associative factors in teachers' ratings. The findings in this study are consistent with their results. It seems clear, however, that teachers' ratings do not reflect the discrete traits of personality as they are presented in a personality inventory. The halo effect which is found so commonly in rating procedure is also difficult to avoid in teachers' ratings.

Ryan has suggested that instead of hoping for literal measures of traits through teacher ratings, they should be utilized in a possible search for other values. The net result may be, as he and other researchers have found, that teacher ratings represent very little more than reflections of academic performance.

Although Barron, Woodworth, and others hold that teachers' ratings are inefficient and unreliable, the findings in this study indicate that they have certain values. The teacher ratings in this study were in accord with the scores of a personality inventory. Whether they can assess specific traits directly, or whether they may predicate their existence on the basis of the known traits is debatable. Nevertheless, their agreement with the standardized inventory is quite evident.

Summary

The significance of teacher ratings has been explored by correlating students' scores on the CPI scales with teachers' ratings of the students on the same scales. For 17 of the 18 scales, the correlation coefficients were significant. The highest correspondence between ratings and scores was shown in intellectual achievement, efficiency, sociability, and responsibility. When considered in reference to findings from other studies, the findings suggest that teacher ratings are most useful in rating personality traits associated with scholastic achievement and sociability.

REFERENCES

1. Barron, F. "Originality in Relation to Personality and Intellect," Journal of Personnel, XXV (1959), pp. 730-42.
2. Benson, G. P. "An Evaluation of Teacher Ability to Identify Maladjustment," Dissertation Abstracts, XVII (1957), pp. 1946-47.
3. Cattell, R. B., and Drendahl, J. C. "Personality and Creativity in Artists and Writers," Journal of Clinical Psychology, XIV (1958), pp. 107-11.
4. Cronbach, L. J. "California Psychological Inventory," in O. K. Buros (ed.), The Fifth Mental Measurements Yearbook (Highland Park, New Jersey: Griphon, 1959).
5. Gough, H. G. California Psychological Inventory Manual (Palo Alto, California: Consulting Psychologists Press, 1957).
6. Holland, J. L. "Some Limitations of Teacher Ratings as Predictors of Creativity," Journal of Educational Psychology, L (1959), pp. 219-23.
7. Liddle, G. "The California Psychological Inventory and Certain Social and Personal Fac-

TABLE II

THE RELATION OF TEACHER RATINGS TO STUDENT SCORES ON CPI SCALES

	Variable	r
Do	Dominance	32**
Cs	Capacity for Status	38**
Sy	Sociability	34**
Sp	Social Presence	25**
Sa	Self-Acceptance	24**
Wb	Sense of Sell-Being	29**
Re	Responsibility	42**
So	Socialization	44**
Sc	Self-Control	26**
To	Tolerance	37**
Gi	Good Impression	18*
Cm	Communality	34**
Ac	Achievement via Conformance	38**
Ai	Achievement via Independence	40**
Ie	Intellectual Efficiency	43**
Py	Psychological-Mindedness	21**
Fx	Flexibility	01
Fe	Femininity	36**

Note--r .05 = 14. r .01 = 18+

*r = .05 level of confidence **r = .01 level of confidence

tors," *Journal of Educational Psychology*, XLIX (1959), pp. 144-48.

8. Leton, D. and Walter, S. *A Factor Analysis of the California Psychological Inventory and Minnesota Counseling Inventory*, report in preparation.

9. Mitchell, J. V., and Pierce, J. J. "A Factor Analysis of Gough's California Psychological Inventory," *Journal of Consulting Psychology*, XXIV (1960), pp. 453-56.

9a. Nunnaly, J. C., and Husek, T. R. "Semantic Clarity: One Standard for Factored Tests," *Educational and Psychological Measurement*, XVIII (1958), pp. 761-68.

10. Ryan, F. J. "Trait Ratings of High-School Students by Teachers," *Journal of Educational Psychology*, XLIX (1958), pp. 124-28.

11. Shaffer, L. F. "California Psychological Inventory," in O. K. Buros (ed.), *The Fifth Mental Measurements Yearbook* (Highland Park, New Jersey: Griphon, 1959).

12. Swab, J. C. and Peters, C. C. "The Reliability and Validity of Estimates as Measuring Tools," *Journal of Educational Sociology*, VII (1933), pp. 224-32.

12a. Tallent, N. "Behavioral Control and Intellectual Achievement of Secondary School Boys," *Journal of Educational Psychology*, XLVII (1956), pp. 490-503.

13. Thorndike, R. L. "California Psychological Inventory," in O. K. Buros (ed.), *The Fifth Mental Measurements Yearbook* (Highland Park, New Jersey: Griphon, 1959).

14. Woodworth, D. G. *A Factorial Study of Trait-Ranking Used in an Assessment of Professional Research Scientists*, paper read at Western Psychological Association, Monterey, California (1958).

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 3, November 1962)

An Investigation of the Relationship of Some Pre-Adolescent Developmental Factors to Adolescent Social Adustment*

ISRAEL WORONOFF

Eastern Michigan University

THE UNIVERSITY of Michigan Child Development Laboratories have been attempting to develop an understanding of the "whole child" through the collection and maintenance of continuous-growth records of the pupils enrolled in the University School.

At the time this study was undertaken, the organismic data of the Child Development Laboratories comprised material only through the elementary grades. Therefore it was felt than an exploratory investigation of the relationship between an individual's developmental factors as recorded in the first through the sixth grades and his adolescent social adjustment would be of value. Through this study, the worth of the longitudinal measures which comprise the Olson-Hughes organismic age variables as predictors of social adjustment would be put to a test. The organismic age is based on the measurement of seven age-unit variables of height, weight, carpal development, dental development, grip strength, mental age, and reading achievement.

Problem

The problem consisted of the examination of the following hypotheses:

1. Slower-growing adolescents have different types of adjustment problems from more rapidly growing adolescents of the same sex.
2. Slower-growing adolescents have a different number of social adjustment problems from more rapidly growing adolescents of the same sex.
3. Slower-growing adolescents receive scores on the teacher rating scale of student social adjustments which are significantly different from those received by rapidly-growing adolescents of the same sex.
4. Slower-growing adolescents receive different sociometric ratings of classmates on friendship and leadership from those received by rapidly-growing adolescents of the same sex.

The sample consisted of 27 girls and 15 boys, aged 13 through 18 years of age. All were graduates of the University School, where longitudinal records were maintained on their growth and development for their entire period of attendance at the school.

Social adjustment is defined as the interaction of an individual with other personalities, groups, and culture elements.[1] In other terms, the investigation concentrates on factors involved in the adolescent's relationship to his peers, parents, siblings, and teachers. No attempt is made in this study to examine the deep-set emotional drives of the teenager; instead, only the overt characteristics of interaction are the determinants of the measuring devices utilized.

Pre-adolescent developmental factors are defined as the organismic variables of height, weight, dental, carpal, grip, mental, and reading ages which were recorded periodically for every child in the sample during his attendance at the University Elementary School.

Slow growers and fast growers are defined operationally. Each subject of the study has three growth statuses rather than one. For a male, the first status is determined by subtracting his chronological age from his present growth age as ascertained by his particular organismic-age line extrapolation, described later. The resulting deviation value—the difference between the growth age and the chronological age—is compared with the amounts of deviation of each of the other males in the study, which are arranged in the following manner: all those whose growth ages are larger than their chronological ages are placed above those whose chronological ages are greater than their growth ages. In the former group, the greater the difference between the two ages, the higher the position on the list; while in the latter group, the greater the difference between the two ages, the lower the position on the list. The reason for this arrangement becomes clear when one realizes that a person whose growth is above his chronological age is growing faster than one of whom the reverse situation is true. Thus, if a male subject's deviation value falls in the top half of the sample, he is classified as a fast grower; while if his position is found to be in the lower half, he is categorized as a slow grower.

---

* Footnotes will be found at the end of the article.

The same operation is followed for the classification of the female sample members, with one difference. The female fast grower has a deviation value which is found in the highest third of the sample, while the slow grower's deviation value falls within the lowest third, because of the larger number of girls than boys in the study.

The second growth status is determined in the same manner as the one just discussed, except for the substitution of a physical age for an organismic one. The physical age is determined by extrapolation also. However, this age is based on only five growth measures, eliminating the mental and reading measures, which are included under organismic age.

The third growth status is determined by the subject's slope value "a", which is ascertained for each sample member in the procedure followed to fit a linear equation to his organismic growth data. This value is compared with the slopes of other sample members of his sex, which are arranged in decreasing order. For males, if the slope being considered is large enough to be placed in the top half of the sample, the subject is classified as a fast grower. Conversely, if it is smaller than that of 50 percent of the sample members, he is categorized as a slow grower.

The procedure is the same for the girls. However, a fast-growing girl must have a slope large enough to be placed in the top third of the sample, while the slow-growing girl's slope value must be small enough to be found in the lowest third of the sample.

As a result of these three methods for the classification of slow and fast growers, it is possible for a subject to be categorized as a slow grower on one basis and as a fast grower on another basis.

Procedure

In order to determine the proper level of growth of each subject, the following procedure was employed: The seven organismic growth measures recorded at the chronological ages of 72, 84, 96, 108, 120, 132, and 144 months were placed on cards, each of which contained all growth data for a single subject. Then the organismic-age variables were averaged at each chronological age interval. In other words, the organismic age of each sample member was computed for every year of his attendance at the University Elementary School from the first to the sixth grades.

If a child did not complete six years at the school, only the years of his attendance were used in this procedure. Nevertheless, no child's records were used unless they covered growth data for at least three successive years. Furthermore, the records of many subjects were found to lack either carpal age scores or dental age measures at various chronological intervals. However, no subject lacked more than one age measure for any particular interval. Whenever this was the case, the average of the remaining age variables was computed and recorded as the organismic age for that interval. A physical age was computed in a similar manner by averaging the height, weight, dental carpal , and grip ages of each individual. Following this procedure, the linear equation $y = ax + b$ was determined for each subject's growth record.[2]

Another elementary school measure which was employed in the statistical analysis is the intelligence quotient, based on the Stanford-Binet test, taken at 120 chronological-age months. The reason for using the measure at this period is that all sample members had a mental age recorded for this interval. This provided still another means of categorizing the sample members into relatively inferior or superior positions.

Method of Investigation

To test the hypotheses, three instruments were utilized— the Mooney Problem Check List, a teacher-rating scale of student social adjustment, and a sociometric test for popularity and leadership. The teacher-rating scale evaluated the following factors in social adjustment: popularity,with classmates of same sex, popularity with classmates of opposite sex, leadership ability, courtesy and consideration, honesty, interest in opposite sex, recipiency of teasing, parental domination, general social adjustment, general affectivity, and self confidence.

The purpose of obtaining the sociometric data was to measure the social stimulus value of sample members to their peers. The students in every homeroom class in which there were subjects in the study were asked to write down the names of classmates with whom they would like to "pal around." They then were asked to write the names of those whom they would like to have represent them at a teacher-student conference to change the school program. Thus, peer-group evaluations of a subject's popularity and leadership traits were brought out. A t test on the measures was employed to check significant differences between the slow and rapid growth groups.

Summary of Results

1. The hypothesis that slow growers have different types and numbers of problems from fast growers of the same sex as determined by responses to the Mooney Problem Check List was not adequately supported by the statistical procedures employed.

2. The hypothesis that slow growers will obtain different social adjustment scores from rapid growers of the same sex is given little statistical sup-

port at high levels of confidence. Trends of statistical significance reveal that fast-growing girls (determined by organismic-age line extrapolations and by slope values) are rated better adjusted by teachers than their slower-developing peers on self confidence and general affectivity. A trend at the ten percent statistical level of confidence indicates that fast-growing males (determined by physical-age line extrapolations) are judged superior to the less physically mature boys in leadership, in ability to cope with the social demands of the school and community, and in their over-all social adjustment.

Although the categorization based on intelligence shows no trend for the males, a number of significant trends emerge for the girls classified on this basis. The female high-level intelligence grouping was judged significantly better adjusted than the lower-IQ girls in affectivity, popularity with the opposite sex, and honesty. Moreover, they received mean adjustment scores superior to their peers at the one percent level of significance.

3. The hypothesis that slow growers will receive different sociometric ratings from fast growers of the same sex does not receive statistical support at high levels of confidence.

The IQ classification revealed a significant trend for the girls. At the two percent level, the girls with the higher mental scores were preferred over the lower group as leaders.

Discussion and Conclusions

The small size and variable-age range of the sample impels this investigation to be classified as an exploratory one. The lack of statistical support at high levels of significance of the several hypotheses examined may be partially the result of the nature of the sample, which did not have enough variation to lend itself to clear-cut categorizations into fast- and slow-growth groupings. The extrapolation technique may have been a factor in the obtained low levels of statistical significance.

Two conclusions from this study are ventured:

1. Studies which employ only biogenic variables in examining adjustment relationships lead to oversimplification of the processes of human behavior and therefore may result in erroneous generalizations.

2. There is an interaction of influence between the growth factors and the personal experiences of an organism, which results in the modification of the effects of both forces in impelling the organism to act along certain paths. Therefore the adjustment of an individual cannot be predicted without a complete understanding of the relationship of these influences within his particular life situation.

FOOTNOTES

* This is a report of one aspect of a larger investigation completed to satisfy the dissertation requirements for the Ph. D. at the University of Michigan, 1953. Appreciation is expressed to Dr. W. C. Olson for his guidance in the planning and execution of the research.

1. Henry Pratt Fairchild (ed.). Dictionary of Sociology (New York: Philosophical Library, 1944), pp. 275-6.
2. Charles C. Peters and Walter R. Van Voorhis. Statistical Procedures and Their Mathematical Bases (New York: McGraw-Hill Book Company, 1940), pp. 428-29.

REFERENCES

1. Lindquist, Everet F. Statistical Analysis in Educational Research (Boston: Houghton-Mifflin Company, 1940), pp. 111-266.
2. Mooney, Ross L. "Exploratory Research on Students' Problems," Journal of Educational Research, XXXVII (November 1943), pp. 218-24.
3. Olson, Willard C. Child Development (Boston: D. C. Health and Company, 1949), pp. 124-34.
4. Olson, Willard C., and Hughes, Byron O. "Growth of the Child as a Whole," Chapter XII in Barker, Kounin and Wright, Child Behavior and Development (New York: McGraw-Hill Book Company, 1943), pp. 199-208.

JeR Field News

HIGH-SCHOOL GRADUATES EXCEED DROPOUTS IN MOST AREAS OF STATUS. This is a finding of a study recently completed by Howard E. Henry at the University of Pennsylvania. The purpose of the study was to determine whether dropouts differ in status from graduates 20 years after leaving school when both groups are matched according to scholastic opportunities and scholastic performance while attending school. Twenty-seven dropouts from the ninth grade in four central Delaware high schools were selected on the basis of complete scholastic records and availability for interview. Each dropout was matched with a graduate of the same sex who had been a ninth-grade classmate of the dropout in 1932, 1933, or 1934. Matched dropout and graduate had attended the same school during grades 1 to 9 and had received comparable school marks from the same teachers in grades 5 and 6.

Through an interview technique it was found that the advantages enjoyed by the graduates logically could be associated with the amount of schooling received. The advantages were greatest in the areas of educational opportunities, employment opportunities and security, and service to society through participation in the beneficient organizations.

* * * * *

THE 20th ANNUAL READING INSTITUTE at Temple University will be held January 26, 1963, to consider "Reading in Curriculum Development." Further information may be obtained from Bruce W. Brigham, coordinator of Institute Services, The Reading Clinic, Department of Psychology, Temple University, Philadelphia 22, Pennsylvania.

* * * * *

A DIVISION OF RESEARCH AND FIELD SERVICES AT TEACHERS COLLEGE, COLUMBIA UNIVERSITY, has been instituted. Dr. Arthur W. Foshay, Professor of Education and executive officer of the Horace Mann-Lincoln Institute of School Experimentation, has been appointed as the director. The College's research program in education is believed to be one of the most extensive in the world. More than 100 research and field-service activities, aimed at improving education in related areas, are underway at the college. The cost of these projects is about $3,000,000 a year with a range of $500 to more than $100,000 for individual studies. Between a third to half of the Teachers College faculty of 141 professors are directly involved in research and field service at any given time. The division will coordinate short and long term research service projects, arrange for the pooling of research personnel, and increase the efficiency with which college facilities are used for research.

* * * * *

PERSONS INTERESTED IN RESEARCH IN ENGLISH may submit proposals for research in this area. Colleges, universities, and state departments of education may submit research proposals directly to: Coordinator of Project English, Cooperative Research Branch, U. S. Office of Education, Washington 25, D. C. Local school districts should submit their proposals through their State education departments. Proposals may be submitted at any time. Deadline dates are April 1, September 1, and December 1 of each year. Decisions are reached 60 to 90 days after each of these dates.

* * * * *

THE NATIONAL SCIENCE TEACHERS ASSOCIATION announces a grant of $20,000 from Ford Motor Company for the support of the Association's 1963 Future Scientists of America Awards for Students. A similar grant will be made for support for the 1964 program. The philosophy of this program is to provide motivation and recognition for research-type scientific investigations by junior and senior high-school students. It is felt that promising students should have experiences resembling those of practicing scientists and engineers, and that these experiences should begin at least as early as the seventh grade.

* * * * *

THE DEPARTMENT OF ROMANCE LANGUAGES AND LITERATURE OF THE UNIVERSITY OF WASHINGTON is conducting a research project to find out what it ought to mean to teach a language or literature in its cultural context. Professor Howard Lee Nostrand will be glad to help any students seeking topics for thesis or term papers bearing on the description or the teaching of a culture. Furthermore, he makes an offer to serve as a clearing agent for any student conducting research in this area.

* * * * *

PERSONS INTERESTED IN ART THERAPY may find the Bulletin of Art Therapy of use to them. The interests of this journal, which is devoted to the visual arts as they contribute to human understanding and mental health, cut across disciplinary lines. Art Therapy shares common ground with art education on the one hand, and conventional psychotherapy and psychoanalysis on the other.

* * * * *

Readers are invited to send items of research interest to John Schmid, Professor, College of Education, University of Arkansas, Fayetteville, Arkansas.

Editorial Board

JACOB O. BACH, Head, Educational Research Bureau, Southern Illinois University, Carbondale, Illinois. Critique Editor.
WILLIAM A. BROWNELL, Dean, School of Education, University of California, Berkeley, California. Education Psychology, Learning and Teaching.
ROBERT A. DAVIS, Professor of Educational Research, George Peabody College for Teachers, Nashville, Tennessee. Evaluative Techniques.
GUY D. McGRATH, Dean, College of Education, Arizona State University, Tempe, Arizona. Secondary Education.
HAROLD E. MITZEL, Assistant Dean for Research, Pennsylvania State University, University Park, Pennsylvania.
WILLARD C. OLSON, Dean of College Education, University of Michigan, Ann Arbor, Michigan. Child Development.
JOHN SCHMID, Professor, College of Education, University of Arkansas, Fayetteville, Arkansas. Field News Editor.
WILSON B. THIEDE, Professor of Education, University of Wisconsin, Madison, Wisconsin. Book Review Editor.
ARTHUR E. TRAXLER, Executive Director, Educational Records Bureau, New York 32, New York. Educational Measurements.

Associate Editors

HAROLD M. ANDERSON, Associate Professor of Education, University of Colorado, Boulder, Colorado.
VERNON E. ANDERSON, Dean, College of Education, University of Maryland, College Park, Maryland.
RUDYARD K. BENT, Professor of Education, University of Arkansas, Fayetteville, Arkansas.
EMMETT ALBERT BETTS, Research Professor, University of Miami, Coral Gables, Florida
LAWRENCE P. BLUM, Assistant Director, Student Personnel Services, University of Wisconsin–Milwaukee, Milwaukee 3, Wisconsin.
J. MURRAY LEE, Chairman, Department of Elementary Education, Southern Illinois University, Carbondale, Ill.
D. WELTY LEFEVER, Professor of Education, University of Southern California, Los Angeles, California.
L. JOSEPH LINS, Professor and Coordinator, Institutional Studies, University of Wisconsin, Madison, Wisconsin.
S. E. TORSTEN LUND, Associate Professor of Education, University of California, Berkeley, California.
WILLIAM A. McCALL, Lecturer in Education, University of Miami, Coral Gables, Florida.
DELYTE W. MORRIS, President, Southern Illinois University, Carbondale, Illinois.
VALWORTH PLUMB, Chairman, Division of Education and Psychology, University of Minnesota (Duluth Branch), Duluth, Minnesota.
DAVID G. RYANS, Rand Corporation, Santa Monica, California.
LESTER B. SANDS, Professor of Education, University of California, Santa Barbara, Goleta, California.
MANDEL SHERMAN, President, and Consultant, Emotional Health Association, Los Angeles, California.
MILDRED B. STANTON, Consultant, Special Education, State Department of Education, Hartford 15, Conn.
W. W. THEISEN, Professor of Education, Marquette University, Milwaukee, Wisconsin.
M. R. TRABUE, Professor of Higher Education, University of Kentucky, Lexington, Kentucky.
J. WAYNE WRIGHTSTONE, Director, Bureau of Educational Research, Board of Education of City of New York, Brooklyn, New York.

Editors Emeriti

E. J. ASHBAUGH, Dean Emeritus, School of Education, Miami University, Oxford MR50, Ohio.
HARRY J. BAKER, 19050 Wiltshire, Lathrup Village, Michigan.
FOWLER D. BROOKS, Director Emeritus, Department of Psychology, DePauw University, Greencastle, Ind.
LEO J. BRUECKNER, Professor Emeritus, University of Minnesota, Minneapolis, Minnesota.
B. R. BUCKINGHAM, 33 Chatham Circle, Wellesley Hills, Massachusetts.
EDGAR A. DOLL, Director Emeritus, Devereux Schools, Devon, Pennsylvania.
HARL R. DOUGLASS, Director Emeritus, School of Education, University of Colorado, Boulder, Colorado.
E. DUNCAN GRIZZELL, Dean Emeritus, University of Pennsylvania, 640 Maxwelton Court, Lexington, Ky.
ERNEST HORN, Professor Emeritus of Education, State University of Iowa, W307 East Hall, Iowa City, Iowa.
EARL HUDELSON, Professor Emeritus of Education, West Virginia University, Morgantown, West Virginia.
A. M. JORDAN, Professor Emeritus, University of North Carolina, Chapel Hill, North Carolina.
ALFRED S. LEWERENZ, Director Emeritus of Evaluation and Research, 4107 Holly Knoll Drive, Los Angeles, California.
A. R. MEAD, Director Emeritus of Educational Research, University of Florida, 1719 N. W. 6th Ave., Gainesville, Florida.
M. J. NELSON, Dean Emeritus, Iowa State Teachers College, 6033 West Orange Drive, Glendale, Arizona.
JESSE B. SEARS, Professor Emeritus, Stanford University, Stanford, California.
CARLETON W. WASHBURNE, Chairman Emeritus, Department of Education, Brooklyn College, Brooklyn, New York.

Staff

JEAN WILLIAMS, Supervisor of Editorial Services. CLARENCE A. SCHOENFELD, Editorial Associate. ARNOLD CAUCUTT, Business Manager.

The Journal of Educational Research

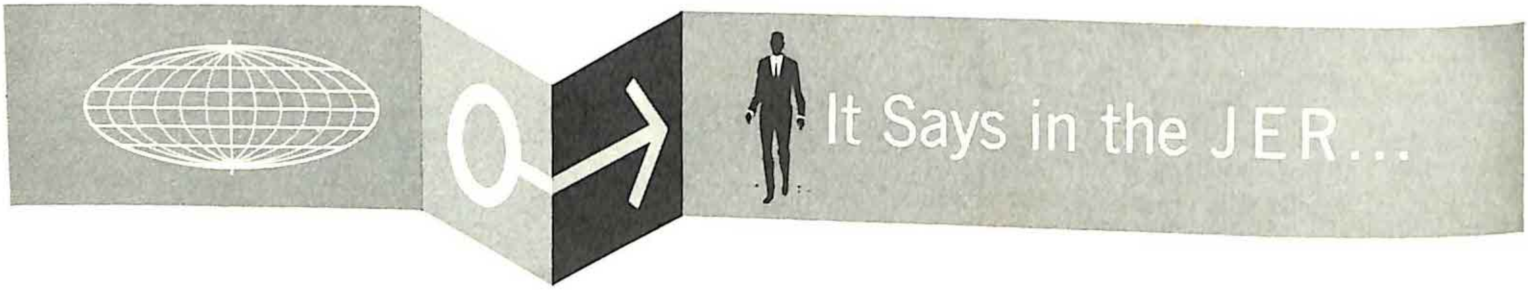
Dedicated to the Scientific Study of Education

VOLUME 56 DECEMBER 1962 NUMBER 4

It Says in The JER 170
In Memoriam 172
Among Those Present 174
Factors Affecting the Academic Underachievement of Bright High-School Students—Helen Erskine Roberts 175
An Investigation of the Differences Between Tenth-, Eleventh-, and Twelfth-Grade "Conforming" and "Nonconforming" Boys—Marvin Powell and Jerry Bergem 184
Comparisons Between Second-, Third-, and Fourth-Grade Children in the Audio-Lingual Learning of Japanese as a Second Language—Robert E. Grinder, Aiko Otomo, and Winifred Toyota 191
A Study of the Effectiveness of Certain Techniques of Reading Instruction in Developing the Ability of Junior High-School Students to Conceptualize Social Studies Content—Mary Witt 198
Readability Versus Reading Ability—Wilbur R. Miller 205
Some Illustrated Comments on Hypothesis Construction and Research—Donald J. Willower 210
Differences in Achievement Test Scores Associated with Item Exposure Time and Audio Reinforcement—R. P. Kropp, F. J. King, W. B. Burr, and J. G. Beard 214
A Formula for Predicting the Comprehension Level of Material to be Presented Orally—John R. Rogers 218
The Effect on Typewriting Performance of Varying the Degree of Interest in Timed-Writing Copy Materials—Willard H. Nelson, Betsy Ann Little, and F. J. King 221
Fields News—John Schmid 223

The Journal of Educational Research is published at Madison, Wisconsin, monthly 10 times a year (the May–June and July–August issues being combined) by Dembar Publications, Box 1605, Madison 3, Wisconsin. The subscription price is $7.50 a year plus $1 a year postage for all subscriptions outside the continental United States. Single copies, $1. Address all communications to Dembar. Second class postage paid at Madison, Wisconsin.

COPYRIGHT 1962 BY DEMBAR PUBLICATIONS, MADISON 3, WISCONSIN



It Says in the JER...

THE CONCEPT APPROACH to the teaching of social studies can be a promising way to help students develop critical thinking, Witt concludes after studying the effectiveness of certain techniques of reading instruction in improving ability of junior-high students to conceptualize social studies content. In related reading and social studies instruction, Witt stressed organizing through outlining and drawing conclusions from factual materials. Her students showed significant gains during each semester in the grasp of 10 selected concepts.

HOW IN FACT DOES A RESEARCHER RESEARCH? Willower suggests we ought to know more about the actual behavior of the social scientist if we are to gain meaningful insights into the practical intricacies of the scientific method. He says creativity of conceptualization plays an integral part, particularly in a study guided by theory. The scientific method is hardly a mechanical process, Willower writes, and researchers are often required to make unanticipated decisions and occasional compromises relative to levels of abstraction and evidence.

$\underline{G = .669\, I + .4981\, LD - 2.0625}$. This is a new formula for predicting the comprehension level of material to be presented orally, proposed by Rogers, in which G is grade level, I is average idea unit length, and LD the average number of words in a hundred word sampling which do not appear on Dale's long list. The -2.0625 is a constant. The formula has not been rigorously tested but would seem to offer an area for fruitful communications research.

WHY DO BRIGHT HIGH SCHOOL STUDENTS fail to earn as good marks in their academic subjects as might be expected of students with their high potential? Roberts suggests that underachievement in school is caused by the student's feelings about thimself and his environment, and that many of the problems related to poor achievement seem to enter school with the child. She says many teachers do not realize that a group test of intelligence does not give an accurate estimate of a bright child's potential, and calls for more individual counseling focused on retrieving the low achiever.

WHAT DISTINGUISHES a disciplinary case from a conforming high school boy? In general, the non-conformer shows poorer social adjustment; he is more impulsive; his reading ability is low and so are his grades; he is often absent or tardy; he holds more part-time jobs, owns his own car, smokes, and stays around "hang-outs." The conforming boys, on the other hand, come from higher-level socio-economic homes, have better educated parents, like school, plan to continue their education, have higher-level vocational goals, and attend church frequently. Powell and Bergem made the study at Euclid, Ohio.

CAN YOU ADMINISTER TESTS BY TELEVISION? Yes, say Kropp, King, Burr and Beard. They presented tests by projecting items onto a screen to be viewed and responded to by students, and got scores similar to those obtained under conventional booklet presentation.

WHAT IS THE OPTIMUM GRADE LEVEL at which to introduce foreign language instruction? Grinder and associates taught children in the second, third, and fourth grades the same audio-lingual curriculum in Japanese for one school year, and found that fourth grade results were much more substantial. The study suggests that foreign language instruction should be introduced later rather than earlier in the elementary school curriculum. The suggestion is contrary to current practice in the United States, but it is in accord with the recommendation of Conant that FL instruction begin in the seventh grade.

IN TYPING INSTRUCTION, does the interest value of the materials to be copied affect typing speed and/or accuracy? Little, Nelson, and King report no significant difference between low-interest and high-interest materials.

TOO DIFFICULT TO READ. That is the indictment of some sections of general shop textbooks written ostensibly for ninth-graders, according to Miller. Until authors and publishers can provide acceptable texts, the industrial arts teacher will have to supply supplementary materials having a lower level of readability, the researcher says.

In Memoriam

MEMORIAL RESOLUTION OF THE FACULTY OF THE UNIVERSITY OF WISCONSIN ON THE DEATH OF PROFESSOR ARVIL S. BARR

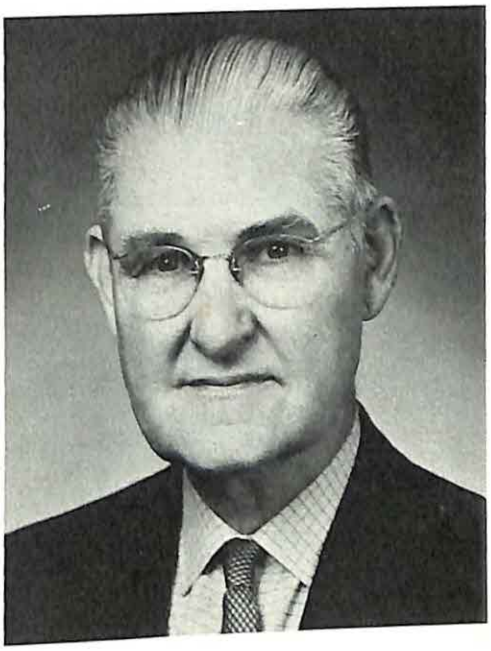
ARVIL S. BARR was born in Selvin, Indiana, January 10, 1892, and died in Madison on May 12, 1962, after thirty-eight years of service at The University of Wisconsin.

He received the degrees of bachelor of arts and master of arts from Indiana University in 1915. He began his teaching career early, serving first as a teacher in a rural school at Selvin at the age of eighteen. Following his graduation from Indiana University, his career in education quickly broadened. He held successive positions as a high-school teacher of history in Yankton, South Dakota, 1915-17; instructor of history in the summer session of Yankton College, 1917; critic teacher in history, Indiana University, 1917-18; during the two-year period 1918-20, he served as high-school secretary of the Young Men's Christian Association of Chicago. From 1920-21, he served as Head of the Department of Education, Evansville College, Indiana, and in 1921-24, he served as director in charge of supervision in the public schools of Detroit. He then came to The University of Wisconsin as associate professor in the Department of Education. He earned the degree of doctorate of philosophy at the University of Wisconsin in 1929, and had been professor of education here since that date.

During his days as a college student at Indiana University, he married Lillian Whittinghill who over the years joined him in providing friendship, hospitality, help, and home to numbers of grateful students at the University of Wisconsin.

Professor Barr's educational interests were broad. His primary teaching interests, however, were in the fields of educational supervision and educational psychology, and his major research interests centered upon the measurement and prediction of teaching efficiency. A highly productive scholar, he published hundreds of articles in journals of psychology and education, and was the author of several widely used textbooks on educational supervision, educational measurement, and educational research. In addition to maintaining a continuous flow of research and writing, Professor Barr found time to manage and edit educational journals. Since 1928 he had been editor of The Journal of Educational Research; since 1932, the editor of The Journal of Experimental Education. Professor Barr took pride in the fact that The Journal of Educational Research was the first journal to be devoted exclusively to the reporting of research in education. In the months before his expected retirement from University duties, he had made plans which would significantly expand the opportunities for educational research at the University of Wisconsin and other institutions. Professor Barr served the larger community of the University on many committees of the faculty and the administration. He also frequently served the state as consultant or participant in educational planning and policy-making.

Professor Barr was widely known in the nation and other parts of the world. He was elected to positions of leadership in the American Educational Research Association, The National Society of College Teachers of Education, and the Society for Pedagogical Studies and Research. The range of his intellectual interests is shown by his memberships and participation in the American Association for Advancement of Science, American Psychological Association, Institute of Mathematical Statistics, Psychometric Society, and the National Education Association.



While his interests and activities ranged far beyond the campus, Professor Barr's heart was close to the institution he served and the students and colleagues who worked with him. His professional career was marked by dedicated teaching, prolific research and writing, and perceptive counselling of a large corps of appreciative students. University records show that more than one hundred students received the doctor of philosophy degree under his direction. Many of these students now have records of significant achievement in the fields of their service. The plans of his former students to honor him near the close of the 1961-62 school year after which he would have retired from University duties were grievously interrupted by his sudden illness and subsequent death.

The professional life of Arvil S. Barr reflected the values he sought to teach. He was curious about educational processes and was actively experimental in seeking answers to educational enigmas. He took great delight in the progress and success of his ventures and those of his students, but he was modest in claim to fame. He was a busy man, but never too busy to counsel his students, to help his colleagues, or to serve his University beyond its normal assignments. He was future-facing. His plans for the years ahead were well-laid and were exciting. During his career-long search to extend knowledge about effective teaching, Professor Barr exemplified one of the most perceptive observations yet known about the good teacher—namely, he was a good man, teaching. His exemplary life and career have added much strength and quality to the purposes and programs of the University of Wisconsin.

MEMORIAL COMMITTEE:

Kai Jensen
Lowell E. Noland
Wilson B. Thiede
John E. Willard
J. Kenneth Little, Chairman

Among Those Present

RUSSELL P. KROPP, F. J. KING, and J. G. BEARD are members of the department of educational research and testing of the Florida State University. W. L. BURR was so affiliated when their manuscript was prepared but subsequently became coordinator of programs and services in the test department of Harcourt, Brace, and World, Inc. Kropp received an EdD from the University of Illinois where he served as research assistant in the Bureau of Educational Research and Service. King received a PhD from the University of Texas where he served as research associate on the Human Talent Project. Burr received an EdD from Florida State University. Previously he had been director of group testing in in the Pinellas County Public Schools (Clearwater-St. Petersburg) Florida. Beard received an MS from Florida State University. He is a doctoral student and is engaged in a project on the development of self-instructional materials on test interpretation.

ROBERT E. GRINDER, now at the University of Wisconsin, was formerly a member of the staff of the department of psychology, University of Hawaii. He has an EdD from Harvard. AIKO OTOMO and WINIFRED TOYOTA are supervisors of student teaching, College of Education, University of Hawaii. Each has an MA from New York University.

WILBUR R. MILLER, assistant professor of Industrial education at Purdue University, is currently working with the professional program of industrial arts teacher education. Prior to his current appointment, Professor Miller received his DE from the University of Missouri which had previously granted his ME degree. Dr. Miller has taught industrial arts in the Laboratory School of Columbia, Missouri, and in the Hazelwood Junior-Senior High School of St. Louis County, Missouri.

MARVIN POWELL is professor of education in the department of graduate programs in education at Northern Illinois University, DeKalb. He holds three degrees from Syracuse. He was formerly on the staff of Western Reserve and of the Willoughby, Ohio, schools. He has published widely in the fields of educational psychology and adolescent psychology. JERRY BERGEM, dean of boys, Euclid High School, Euclid, Ohio, holds degrees from Western Reserve, Cleveland, and Ohio.

JOHN R. ROGERS was born in Texas and received a BA from Sam Houston State Teachers College in Huntsville, a ME from the University of Texas, and a PhD from the same institution. He has taught in the public schools of Texas and in the following institutions: Southern State College, Magnolia, Arkansas; Northwestern State College, Natchitoches, Louisiana; and Fort Hayes Kansas State College. At the present he is associate professor of education and director of student teaching at the University of Mississippi.

BETSY ANN LITTLE is administrative assistant to the president of the Dade County (Florida) Junior College. She received her BS from Nasson College and her MS in higher education from the Florida State University. WILLARD H. NELSON is an associate professor of psychology at the Florida State University. He received his BS from Nebraska. He was a post-doctoral fellow at the University of Illinois, and was previously an associate professor of psychology at Auburn University. F. J. KING is assistant professor in educational research and testing at the Florida State University. He received the PhD in educational psychology from the University of Texas.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 4, December 1962)

Factors Affecting the Academic Underachievement Of Bright High-School Students

HELEN ERSKINE ROBERTS
Santa Ysabel, California

BECAUSE bright students usually score at or above the norms for their ages on standardized achievement tests, teachers are inclined to believe that these students are making satisfactory progress. However, this is often not the case. Learning very early that they can get by with a minimum of effort, bright students frequently do not work to their optimum capacity.

The purpose of the present investigation was to find some answers to the question "Why do bright high school students fail to earn as good marks in their academic subjects as might be expected of students with their high potential?"

Other research studies have indicated that there is no single answer. Eells (1), Holden (7), Miner (9), Roper (11), Shaw (12), and many others have determined that lack of achievement is related to the home background of the students, especially as it concerns the socio-economic status and education of the parents. The conclusions are that under-achievement is a phenomenon not specifically caused by the school situation, and that hostilities not related to school are fixed in the personality of the student, often before he enters the first grade.

The present study was undertaken to find some reasons for the poor achievement of a group of students in one high school. An attempt was made to find some of the underlying causes for the differences in academic achievement of bright students in the same high school who showed by their performance on the Iowa Tests of Educational Development that they were gaining knowledge. The problem was, "Why did these bright students fail to get good marks?" The major hypothesis to be tested was, "Underachievement in school is caused by the student's feelings about himself and his environment."

Procedure

The subjects of the study were two groups of students in a Los Angeles high school. Their intelligence on a standard measure was approximately the same, their scores on the Iowa Test of Educational Development were above the ninetieth percentile, and yet their school marks range from "D" averages to "A" averages. The differences between these high- and low-achieving students were assessed in the following areas: cognitive, home background, physical health, personality, school adjustment, and vocational goals.

The instruments used were the Otis Quick-Scoring Mental Ability Tests, Beta and Gamma Forms; the Iowa Tests of Educational Development, Form Y3; the Stanford-Binet individual tests of intelligence, Form L; a questionnaire designed by the investigator; the health, attendance, and cumulative records kept in the high school; and the California Psychological Inventory.

A survey of the results of the Iowa Tests of Educational Development, given the preceding March, showed that 109 boys (17.1 percent of the total number of boys) and 87 girls (14.6 percent of the total number of girls) had scores above the ninetieth percentile. Of these, 28 boys and 22 girls, designated as high achievers, had grade-point averages above 3.60, on the basis of four points for an "A" to one point for a "D." The grade in physical education was not included. Forty-five boys and 19 girls with averages below 2.60 were designated as low achievers.

During the fall semester, a Stanford-Binet individual test was given to each student. In order to prevent variations in the students' IQ's caused by the personalities of different examiners, the same psychometrist gave the Binet to every student. Rapport in all cases was good or excellent, and the IQ's obtained probably represented a fairly accurate estimate of the potential ability of the students. The Binet IQ's ranged from 125 to 154, in contrast to the Otis IQ's, which ranged from 102 to 135.

The mean difference of the Binet and Otis IQ's was 17.54 for the low-achieving students and 15.56 for the high achievers. A comparison of these differences gave a t-value of only -1.326, significant at the 20 percent level. Thus, although the low achievers made a larger gain on the Binet over the Otis than did the high achievers, this difference was without statistical significance.

After the completion of the Binet testing, the 37 members of each group were selected, matched as nearly as possible for Binet IQ's, age, sex and grade in school. The 25 boys and twelve girls designated as high achievers were selected from the students with grade-point averages above 3.60. The low achievers, selected from students with averages below 2.60, had a mean grade-point average of 2.26. In contrast, the average of the high achievers was 3.85. The difference was significant at the one percent level.

Cognitive Differences

Some differences were found between the two groups in intellectual ability, although the difference on the Stanford-Binet was without significance. The mean Binet IQ of the high achievers was 134.76, while that of the low achievers was 133.59. The means of the Otis IQ's were 119.19 for the high group and 116.05 for the low, with the difference being significant at the .01 level.

A comparison was made of the Iowa Tests scores of the high and low achievers. As shown in Table I, the most striking difference appeared to be the superiority of the high achievers on the test of correctness in writing (III) and a difference favoring the low achievers on the test of interpreting literature (VII). These differences may be related to personality factors, as Haggard suggested (6:396), rather than to purely cognitive differences. This difference between the two groups on the test of correctness in writing may provide a clue to the significant difference in their school achievement. The low achievers evidently had paid little attention to detail or rules during their entire school career. On the other hand, they were able to interpret literature more successfully than the high achievers, and on three other tests, the two science and the general vocabulary, their scores were almost as high as those of the high achievers.

Differences in Home Backgrounds

There were a few significant differences in the home backgrounds of the students. The fathers of the high achievers were engaged in higher-ranking occupations than were the fathers of the low achievers. This is in accord with the findings of Terman (16), Frankel (2), Klausmeier (8), Slocum (14), Sutcliffe (15), and Westfall (17). The mothers of the majority of the students of both groups were housewives. Thirteen mothers of high achievers and 14 mothers of low achievers worked outside the home. Five high-achieving students and five low achievers came from homes broken by death or divorce.

The fathers of the high achievers had had slightly more formal education than had the fathers of the low achievers, as shown by the means of 13.56 and 13.16 years, respectively. For the mothers, too, the same difference appeared; the mean number of years of education for the mothers of the high achievers was 13.02, and for the mothers of the low achievers, 12.48. The fathers of the high achievers were significantly younger (.01 level) than those of the low achievers. The mothers of the high achievers also tended to be younger, although the difference was less marked (.07 level).

The chi-square test was used to determine the significance of the number of only children in each group, as well as the number of youngest and oldest children. The value for "only child" was 0.1814, indicating that there was no significant difference between the high and low groups on this variable. There were more oldest children among the high achievers. The value for "oldest child" was 1.95, significant at the .17 level of confidence. However, there were more youngest children among the low achievers than among the high achievers. The chi-square value of 3.01 is significant at the .08 level. These findings suggested that oldest children might have developed a stronger sense of responsibility than had youngest children.

No significant differences were found in the size of the families, the number of home duties performed, nor the age at which home duties were begun.

Concerning the number of high and low achievers who were Jewish, there was a slight difference in favor of the high schievers, among whom there were ten Jewish students, in contrast to three among the low achievers. The chi-square value for this difference was 3.36, significant at the .07 level.

A very significant difference appeared in the amount of time students spent watching television in their homes. The high achievers spent, on the average, 4.43 hours a week, while the low achievers spent 9.95 hours. This difference was significant at the .01 level.

In contrast to this time spent by the low achievers in passive entertainment, many more high achievers indicated that they were taking or had taken music lessons, the chi-square value being significant at the .01 level.

The comments written by the elementary school teachers on the school record cares were not quantified, but they shed some light on the difference in the home backgrounds of the high and low achievers. On the records of the low achievers, there were frequent comments indicating either a disturbed home or a poor relationship between the parents and child. The comments about the parents and homes of the high achievers were usually complimentary.

Differences in Physical Health

The fact that the low achievers were absent more than twice as much as the high achievers suggested that the high achievers might be healthier. However, in spite of the great discrepancy in the at-

TABLE I

ANALYSIS OF IOWA TEST RESULTS

Test	Means		t-Value	Significance
	Highs	Lows		
1. Social Concepts	94.00	91.03	2.22	.05
2. Natural Science	91.92	91.68	.12	---
3. Correctness of Expression	93.78	83.51	4.52	.01
4. Quantitative Thinking	95.81	89.97	3.67	.01
5. Reading in Social Studies	95.46	91.27	3.30	.01
6. Reading in Science	94.75	93.35	1.41	---
7. Interpretation of Literature	86.89	90.10	-1.23	---
8. General Vocabulary	92.35	92.22	0.097	---
Composite Score	96.70	94.76	3.53	.01

TABLE II

HEALTH FACTORS CONSIDERED

Factor	Highs	Lows	Chi-square	Significance
Allergies	10	15	0.9665	.34
Broken Bones	10	15	0.9665	.34
Operations	18	25	1.9985	.16
Asthma	1	7	3.5038	.065
Severe Acne	0	4	2.3786	.13
Hay Fever	7	5	0.0995	---
Daily Breakfast	34	31	0.5060	---
Use of Vitamins	18	11	2.04	.16

tendance records of the high and low achievers, the school health cards indicated little difference in the health of the students. Nor did the students differ in rating their own health. The fact that the members of both groups had had the same number of children's diseases was further indication of the homogeneity of the groups in physical health. Furthermore, in each group, the same number of boys had been approved for the sports program of the school. The number of students wearing glasses was equal, and the mean number of hours of sleep was the same. However, there were slight differences between the high and low achievers on several of the factors considered. In all but one case, hay fever, the differences favored the high achievers, as shown in Table II.

The comments written by the elementary teachers on the school record cards were of such nature that they could be quantified. Only six high achievers were perceived as having health problems, in contrast to 18 low achievers. The chi-square value for this difference is significant beyond the .01 level.

On the basis of these data, it is apparent that slight differences in physical health may have contributed to the success of the high achievers and the failure of the low achievers in their school work.

Personality Differences

Since the major hypothesis to be tested in this investigation was "Underachievement in school is caused by the child's feelings about himself and his environment," it was necessary to select a personality test that would measure numerous characteristics of normal students. The self-administering California Psychological Inventory, with its 18 scales, was given at the high school under the supervision of the investigator. Since the personality patterns for girls and boys may be expected to differ, the data were analyzed separately.

Analysis of Girls' Responses

Even before the statistical computations were made, it was clear that the high-achieving girls had profiles very different from those of the low achievers. The scores of the high achievers usually fell above the norm for high-school females, and the scores formed a consistent pattern. In contrast, great internal variability was noted among the scores of the low achievers. Statistical tabulation confirmed these differences, as shown in Table III.

The mean scores of the high-achieving girls were higher than those of the low achievers for all scales except Flexibility. Pierce (10) also found the low-achieving girls in his study more flexible in their attitudes than the high achievers.

In the present investigation, the differences were significant at the .04 level in Social Presence and at the .01 level in Dominance, Sociability, Well-being, Responsibility, and Achievement via conmance. On the Socialization and Self-Control Scales, the differences favored the high achievers at the .07 and .08 levels.

The differences between the two groups of girls, which showed clearly in their responses to the questions on the California Psychological Inventory, had been apparent when they were in elementary school. Some first-grade teachers had commented that the high achievers were dependable, well-adjusted, poised, and conscientious about their work. The comment of one teacher was "An exceptional child. Loves to do everything she is supposed to do and does it all well." In contrast, the teachers' comments about the low achievers indicated personality problems as early as the first grade for some of the girls. About one such girl, the first-grade teacher had written "Unstable, excitable, nervous, able to do good work if held down."

The conclusion to be drawn is that although the low-achieving girls had the intellectual ability to get good marks in school, problems associated with their personality development prevented their doing so. Thus, for the girls, the hypothesis that underachievement in school is related to a child's feelings about himself and his environment was substantiated.

Analysis of Boys' Responses

As was the case with the girls, the individual profiles of the high- and low-achieving boys were very different. The boys who had been successful in school tended to have regular profiles showing less internal variability than did those of the low achievers. The statistical computations, given in Table IV, show that there were important differences between the two groups.

The most significant difference (.01 level) between the scores of the two groups appeared on the Socialization Scale, which is in agreement with the findings of Gough (4, 5) and Pierce (10). The difference on the Femininity Scale was at the .02 level and at the .04 level on Achievement via conformance. The fact that the groups had the same mean score on Intellectual efficiency indicates that both the high and low achievers were intelligent, alert, and well-informed. As described on Gough's scale of Socialization, the high achiever differs from the low in being serious, honest, industrious, modest, obliging, sincere, and steady, rather than defensive, demanding, resentful, stubborn, headstrong, rebellious, and undependable (3).

The differences between the two groups of boys on the scales of responsibility, self-control, and good impression, although less significant, reinforce the picture of the high achiever as an individual who is responsible, conscientious, cooperative, diligent, and persistent, in contrast to the low achiever, who tend to be impulsive, self-centered, uninhibited, and little concerned with the

TABLE III

SCORES OF GIRLS ON CALIFORNIA PSYCHOLOGICAL INVENTORY

Scale	Mean	S.D.	Range	t-Value	Significance
Do - H	30.25	5.10	20-37	4.201	.01
- L	22.50	3.38	15-27		
Cs - H	21.08	3.52	14-27	1.17	.25
- L	19.08	4.43	12-27		
Sy - H	25.50	4.96	15-31	2.82	.01
- L	19.83	4.43	14-27		
Sp - H	36.33	6.30	27-47	2.18	.04
- L	31.00	5.10	26-42		
Sa - H	21.33	2.72	16-26	0.956	.34
- L	20.17	2.99	15-26		
Wb - H	36.75	2.68	31-41	3.326	.01
- L	32.58	1.00	27-38		
Re - H	33.75	3.14	27-39	2.78	.012
- L	28.83	3.46	23- 5		
So - H	40.67	5.95	26-48	1.875	.074
- L	35.33	7.32	21-48		
Sc - H	28.58	7.07	19-39	1.777	.088
- L	24.08	4.54	15-29		
To - H	24.75	4.09	13-30	0.984	.338
- L	23.17	3.43	17-30		
Gi - H	16.42	5.65	5-25	1.352	.19
- L	13.58	4.05	8-22		
Cm - H	26.42	-.85	25-28	0.823	.425
- L	25.83	2.20	20-28		
Ac - H	27.67	3.30	21-33	2.756	.012
- L	23.17	4.30	16-29		
Ai - H	20.33	2.50	16-26	-.090	.80
- L	20.25	2.35	16-23		
Ie - H	41.33	4.61	32-48	1.595	.128
- L	38.58	3.38	34-46		
Py - H	11.08	2.57	7-16	1.071	.298
- L	10.00	2.16	7-13		
Fx - H	11.75	4.21	6-18	-0.984	.338
- L	13.25	2.80	9-20		
Fe - H	25.33	3.47	20-30	1.203	.245
- L	23.67	3.01	17-27		

TABLE IV

SCORES OF BOYS ON CALIFORNIA PSYCHOLOLOGICAL INVENTORY

Scale	Mean	S. D.	Range	t-Value	Significance
Do - H	27.12	6.38	12-36	-0.044	---
- L	27.20	6.16	17-39		
Cs - H	19.60	3.67	12-26	0.921	.365
- L	18.72	3.90	13-27		
Sy - H	25.00	4.77	16-32	0.474	.60
- L	24.36	4.59	16-33		
Sp - H	34.32	4.68	26-45	-0.738	.468
- L	35.36	5.08	26-45		
Sa - H	20.96	3.67	13-27	-0.687	.50
- L	21.04	4.37	13-29		
Wb - H	35.20	4.04	28-41	-0.250	.80
- L	35.48	3.71	28-42		
Re - H	32.24	3.50	24-38	1.285	.202
- L	30.72	4.62	22-41		
So - H	41.00	4.86	31-46	2.689	.01
- L	36.76	6.01	25-48		
Sc - H	28.84	6.01	15-39	1.391	.168
- L	26.20	7.09	14-41		
To - H	22.68	3.55	18-29	0.323	.70
- L	22.32	4.15	14-28		
Gi - H	18.32	6.66	8-31	1.494	.143
- L	15.56	6.13	5-28		
Cm - H	26.12	1.51	23-28	0.508	.60
- L	25.88	1.75	20-28		
Ac - H	28.08	4.23	18-35	2.12	.039
- L	25.28	4.89	16-34		
Ai - H	19.68	4.04	11-28	-0.107	.90
- L	19.80	3.70	10-27		
Ie - H	40.00	4.13	32-46	0.00	---
- L	40.00	4.32	29-48		
Py - H	11.20	2.50	7.17	-0.185	.80
- L	11.32	1.95	8-16		
Fx - H	9.68	4.42	1-21	-0.610	.50
- L	10.44	4.22	1-18		
Fe - H	17.44	2.97	12-22	2.348	.023
- L	15.20	3.61	9-21		

TABLE V

ANALYSIS OF RESPONSES TO SIXTEEN QUESTIONS OF SOCIALIZATION SCALE

	Highs	Lows	Chi-square	Significance
62. I think Lincoln was greater than Washington.	20	15	1.523	.22
156. I hardly ever get excited or thrilled.	1	5	1.705	.19
164. My parents have often disapproved of my friends.	3	7	1.125	.29
168. My home life was always happy.	18	12	2.083	.14
214. In school I was sometimes sent to the principal for cutting up.	3	5	0.149	.70
257. I often feel as though I have done something wrong or wicked.	9*	5	0.893	.35
336. Sometimes I used to feel that I would like to leave home.	8	15	2.90	.09
367. My home life was always pleasant.	16	10	2.00	.15
369. I seem to do things I regret more often than other people.	8*	2	3.125	.08
389. I get pretty discouraged when a smart lawyer gets a criminal free.	20	16	0.893	.35
394. Even when I have gotten into trouble I was usually trying to do the right thing.	18	12	2.083	.14
396. I sometimes wanted to run away from home.	7	11	0.781	.39
431. As a youngster in school I used to cause the teachers lots of trouble.	2	6	1.495	.22
436. I never cared much for school.	0	4	2.45	.12
439. The members of my family were always very close to each other	18	14	0.781	.39
457. A person is better off if he doesn't trust anyone.	1	3	0.272	.61

* The answer "true" given by the high achievers is considered a wrong answer on these questions.

needs and wants of others.

Since the most significant difference between the boys who earned good marks in school and those who had very low marks was found in their replies to the questions of Gough's Socialization Scale, an analysis was made of the answers to the 54 questions of this scale. On the individual questions, there was no difference significant at the .01 to .05 levels, as shown in Table V, in which are presented the 16 questions on which slight differences occurred.

On the Self-acceptance Scale, the scores of the low-achieving boys fell into two groups. Sixteen boys had very high scores, and the remaining nine had very low scores on this scale. In the light of Shippee-Blum's findings (13:30) that the rebellious students in her investigation had extreme scores, either high or low, on the measure of self-esteem, it is noteworthy that the same condition arose for the low achievers in the present study who had been identified not because of other forms of rebelliousness but because of having earned poor marks.

Study of the comments of the elementary teachers showed differences between the high- and low-achieving boys similar to those recorded for the girls. It was clear that the elementary teachers had perceived the two groups of boys as different kinds of students many years before they reached high school. This is in agreement with Terman's findings that parents and teachers were able to discern personality differences between his successful and less successful males 18 years prior to their classification on the basis of their adult achievement (16: 350).

Although the personality differences between the two groups of boys were not so striking as those between the two groups of girls, it is possible to state that the findings reported substantiate the hypothesis that underachievement is closely related to the child's feelings about himself and his environment.

Differences in School Adjustment

The school adjustment of the high and low achievers differed in many ways in addition to the significant difference (.01 level) in the marks they had earned in their academic subjects. The high achievers spent more hours studying (significant at the .01 level), and spent less time looking at television, reading, or attending moving-picture theaters. They participated in more school activities and held more school and club offices. The difference for both was at the .01 level. Even in physical education, the high achievers were more successful than the low achievers, the difference being significant at the .01 level.

Only two high achievers had repeated a subject or grade in school, but 18 low achievers had done this. No high achievers were considered "discipline problems" by the vice-principals, while four low achievers were so designated. The difference between the numbers of absences for the two groups was significant at the .01 level, the lows having been absent more than twice as much as the high achievers.

The poor school adjustment of the low achievers seemed to be a further manifestation of their inability to adjust to their environment.

Differences in Vocational Aspirations

There were very few differences in the vocational aspirations of the high and low achievers. Every student in both groups indicated that he was making his choice of a vocation on his own responsibility. All hoped to attend college and to become professional men or women. However, there was a difference in the types of colleges listed by the members of the two groups. Almost all of the high achievers gave the names of one or more well-known, four-year colleges with strict entrance requirements. In contrast, the majority of the low achievers said that they expected to have to attend a junior college. Another difference was that the low achievers frequently listed a profession as their first choice for a future occupation, and then for a second choice gave a job at a lower occupational level. Finally, every high achiever indicated that he would try to obtain a scholarship. Twenty-one low achievers also wrote that they expected to try for a scholarship, although, because of their school marks, there was little hope that they might be able to get one.

Implications for Practice

The implications for practice deriving from this investigation fall into two categories: those related to the findings of the difference between Otis and Stanford-Binet IQ's and those related to the findings concerning the many differences between high and low achievers.

It is possible that many teachers do not realize that a group test of intelligence does not give an accurate estimate of a bright child's potential. This is especially true for very gifted children, under-achieving children, and children from homes of low socio-economic status. Thus, efforts should be made to inform all teachers about the differences in intelligence quotients derived from various group and individual intelligence tests.

Both elementary and secondary schools should attempt to employ more trained counselors who can test children individually, counsel these children, and confer with their parents. The counselor could do much to prevent or correct many of the problems that confront the low achiever. If there could be an adequate staff of trained counselors in every school, a counselor could work with each child from the moment he enters school and could thus prevent much low achievement, since many of the problems related to poor achievement seem to enter school with the child.

As Terman pointed out, intellect and achievement are far from perfectly correlated. Much remains to be discovered as to why some brilliant students fail to use their intelligence. Today the world needs the contribution of every gifted child, not only in the realm of scientific and mathematical research, but far more in the realm of statesmanship. To find ways of enabling each gifted underachiever to make his contribution to society is the task of all educators.

REFERENCES

1. Eells, Kenneth, et al. Intelligence and Cultural Differences--A Study of Cultural Learning and Problem Solving (Chicago: University of Chicago Press, 1952).
2. Frankel, Edward. "A Comparative Study of Achieving and Under-achieving High School Boys of High Intellectual Ability," Journal of Educational Research, LIII (January 1960) pp. 172-80.
3. Gough, Harrison G. California Psychological Inventory (Palo Alto, California: Consulting Psychologists Press, Inc., 1957).
4. Gough, Harrison G. "Factors Relating to the Academic Achievement of High School Students," Journal of Educational Psychology XL (February 1949), pp. 65-78.
5. Gough, Harrison G. "What Determines the Academic Achievement of High School Students," Journal of Educational Research (January 1953), pp. 321-31.
6. Haggard, Ernest A. "Socialization, Personality, and Academic Achievement in Gifted Children," School Review (Winter 1957), pp. 388-414.
7. Holden, Raymond H., and Isaksen, Henry L. "Fathers' Occupations and Intelligence Level of a Homogeneous Occupational Group," The Personnel and Guidance Journal, XXXVIII (October 1959), pp. 125-27.
8. Klausmeier, Herbert J. "Physical, Behavioral, and Other Characteristics of High- and Lower-Achieving Children in Favored Environments," Journal of Educational Research, LI (April 1958), pp. 573-81.
9. Miner, John B. Intelligence in the United States (New York: Springer Publishing Co., Inc., 1957).
10. Pierce, James V. The Educational Motivation Patterns of Superior Students Who Do and Do Not Achieve in High School, Research Project 208 (7136) U. S. Office of Education, Department of Health, Education, and Welfare(1959).
11. Roper, Elmo. Factors Affecting the Admission of High School Seniors to College (Washington, D. C.: American Council on Education, 1949).
12. Shaw, Merville C., and Grubb, James. "Hostility and Able High School Under-Achievers," Journal of Counseling Psychology (Winter 1958), pp. 263-66.
13. Shippee-Blum, Eva Maria. "The Young Rebel: Self-regard and Ego-Ideal," Journal of Consulting Psychology, XXIII (February 1959) pp. 44-50.
14. Slocum, W. L. "Educational Planning by High School Seniors," Journal of Educational Research, LI (April 1958), pp. 583-90.
15. Sutcliffe, Charles Ernest. Factors Related to Low Achievement by High School Pupils of High Mental Ability, unpublished doctoral dissertation, University of Southern California (1959).
16. Terman, Lewis M., and Oden, Melita H. "The Gifted Child Grows Up," Genetic Studies of Genius, Vol. IV (Stanford, California: Stanford University Press, 1947).
17. Westfall, Frank Wilbur. Selected Variables in the Achievement or Non-achievement of the Academically Talented High School Student, unpublished doctoral dissertation, University of Southern California (1958).

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 4, December 1962)

An Investigation of the Differences Between Tenth-, Eleventh-, and Twelfth-Grade "Conforming" and "Nonconforming" Boys

MARVIN POWELL
Northern Illinois University

JERRY BERGEM
Euclid High School

THE PURPOSE of this study was to determine the differences between "conforming" and "non-conforming tenth-, eleventh, - and twelfth-grade boys in school grades, school attendance, reading personality, temperament, and home environmental conditions.

Review of the Literature

Ruth Strang (14) states that "from 1885 to 1937, there were few references to objective studies in the published articles on discipline". Many "experts" have written articles for popular magazines, but seldom for professional or research journals.

Douglas (3), Krantz (7), and Owens (12) have discussed the administrative procedures related to discipline in the public schools, but have not been overly concerned with the effect of discipline on the adjustment of the disciplined pupil.

Most of the investigations— McClure (9), Boynton and McGraw (1). and Coleman (2), have indicated that far more boys than girls are considered behavior problems. Mullen (10) has shown a ratio of seven boys to three girls in his studies of truancy. Eaton (4) stated, "The data clearly reveal not only that there was more problem behavior among boys, but also that there were more boys involved in almost all types of behavior problems." Olson (11) and Snyder (13) also confirm these findings.

A few studies have attempted to find differences between children who need discipline and those who do not. Snyder (13), Hales (6), Yourman (17), and Tallent (15) have studied differences in intelligence, generally finding that discipline problems show lower average IQ's than non-discipline problems. Tucker (16) found some personality differences based upon standardized tests and stated that "in general, the personality traits of the troublesome children are less desirable than those of the average child of his age".

Socio-economic background has also been studied by Owens (12), Lewis (8), and Snyder (13). The findings are generally inconclusive, but do indicate that more behavior problems come from lower socio-economic backgrounds. Tucker (16) found a relationship between parents' lack of education and behavior.

The present study attempts to expand some of the areas previously investigated and to investigate other variables not previously studied.

Procedure and Subjects

In the large metropolitan high school under study*, a record is kept of every student who is sent to the office because of misbehavior in the classroom. Reasons for being sent include creating a disturbance in class, being constantly tardy to class, fighting, defying the teachers, cutting class, etc. The records for boys are more complete than those for girls, and are kept more up-to-date. A "conforming" student was identified as one who has never been sent to the office (i.e., Dean of Boys, Principal, Assistant Principal) for discipline, and is generally well-behaved in school. A "non-conforming" student is one who has been sent to the office for discipline three or more times during the school year, and is generally considered to be poorly behaved.

Over-all school achievement was determined by averaging the student's grades in his subject-matter courses for the entire year studied. Reading ability was determined by a standardized test, as were personality and temperament. The student's home background was evaluated from data on the school's permanent record card and from data obtained through use of a questionnaire designed by one of the authors.

Approximately 50 male "conforming" and 50 male "non-conforming" students from each of the three grades studied— ten, eleven, and twelve— were used as the control and experimental groups respectively. The groups were matched according to IQ, as determined by the California Test of Mental Maturities. The previous year's records were used to determine in which group a boy belonged. Thus, a tenth grader was considered on the basis of his ninth-grade record, an eleventh grader on his tenth-grade record, and a twelfth grader on his eleventh-grade record.

---

* The authors wish to thank Mr. Clifford Owen, principal of Euclid Senior High School, Euclid, Ohio, for his kind cooperation.

Fifty students were tested at a sitting during a two-hour interval. Those who were absent were given make-up tests. The following tests were administered: 1) the California Test of Personality, 2) the California Reading Test, 3) the Thurstone Temperament Schedule, and 4) the author-designed questionnaire. The first three are accepted, commonly-used tests which the authors felt would adequately measure the variables involved in this investigation. The raw scores on the California tests and the Thurstone Temperament Schedule were used for grades and the questionnaire results. Total grade-point average was a major variable, but an analysis was also done of English and physical education grades. These latter two areas are required of all students, whereas the other courses making up the total grade-point average vary greatly.

The author-designed questionnaire was concerned with the following questions: age at the time of the test, with whom the student lives, the size of the family, whether the student is the oldest or the youngest in the family, the ocuupation of the mother and the father, the education of the mother and the father, the mobility of the family, whether the student has a part-time job, whether the student has a car, whether the student has had any dates, how much he studies outside school, the number of extracurricular activities, time he spends with his parents, how much the student feels he argues with his parents, whether the student smokes, his personal view on how he is disciplined at home, his attitude toward the school, his attitude toward other students, his attitude toward the teachers, his attitude toward other adults, the number of friends he has, whether he desires to finish high school and go on to college, the curriculum he is pursuing, his vocational choice, how he spends his leisure time, the frequency of church attendance.

The permanent office record cards were used to determine attendance, grades, tardiness, etc. Grades were recorded in percentages and/or point averages, and these were tabulated directly.

All differences were studied statistically in terms of the null hypothesis. The t-test was employed to test the significance of the differences. The formula employed (5) to determine the significance of the differences was:

$$\frac{t = M_1 \text{------} M_2}{\left(\frac{d^2 1 + d^2}{(N_1 + N_2 - 2)}\right)\left(\frac{N_1 + N_2}{(N_1\ N_2)}\right)}$$

The level of confidence was checked and all differences significant at the .05 level of confidence or better were considered as meaningful.

The questionnaire data were converted to percentages and presented in tabular form. No precise statistics were applied to these data and the discussion is based on a qualitative analysis of the material.

Results and Discussion

An analysis was made to determine the differences between the tenth-, eleventh-, and twelfth-grade "conforming" and "non-conforming" boys on the variables studied. Table I presents the differences between the two groups on the various sections of the California Test of Personality. The "conforming" tenth-grade boys demonstrated significantly better adjustment (.01 level of confidence) in Social and Total Adjustment, but there was no significant difference between the groups in Personal Adjustment. The same findings appeared at the eleventh-grade level where these differences again were significant (.05 level of confidence), showing better adjustment for Social and Total Adjustment for the "conforming" boys. Again, there were no differences between the groups on Personal Adjustment. There were no significant differences at the tenth-grade level. It would appear that the Social Adjustment difference is the major factor, since it does enter into and may be responsible for the difference in Total Adjustment. It is possible that this failure to make a good adjustment to the other students is one of the factors leading to non-conforming behavior.

Table II presents the data showing the difference between the groups in the various areas of the Thurstone Temperament Schedule. There were no significant differences between the groups at the tenth-grade level. The "non-conforming" eleventh-grade boys appeared to be significantly (.05 level of confidence) more active, vigorous, impulsive, while the "non-conforming" twelfth-grade boys were significantly more impulsive (.05 level of confidence). Impulsiveness does appear to be the sort of trait that could easily get a student into difficulty.

Table III presents the differences between the groups on the various sections of the California Reading Test. The "conforming" tenth-grade boys are significantly better (.01 level of confidence) than the "non-conforming" on both parts of the test as well as the total. At the eleventh-grade level, the difference was only significant for Reading Comprehension (.05 level of confidence), while at the twelfth-grade level, there were significant differences (.05 level of confidence) for the Reading Comprehension and the total test results. Again, the differences favor the "conforming" groups. The reading difficulty indicated here may be either a cause of non-conformity or a result of it. That is, the student may not conform because of his inability to keep up with the others, or he may be a poor reader because of his other adjustment problems.

Table IV presents the differences between the groups in their total grade average, English grade, and physical education grade. In all cases, the differences were significant at the . 01 level of confidence, with the "conforming" boys being the better students. Grades are, of course, often the subjective judgment of the teacher, and low grades may be a reflection of the student's failure to conform. On the other hand, the student may become a non-conformist because his work is poor in relation to the work of his peers.

Table V presents the differences between the groups in relation to the number of days absent and the number of times they were tardy. At the tenth- and eleventh-grade levels, the "non-conforming" group was absent more often (. 01 level of confidence) than the "conforming" group. At the eleventh- and twelfth-grade levels, the "non-conformers" were tardy more often (.01 level of confidence). Both absence and tardiness are probably further indicators of non-conforming behavior. These factors often appear in studies of juvenile delinquents.

Results of the Qualitative Analysis of the Questionnaire

The questionnaire data were qualitatively analyzed and it was found that more tenth-, eleventh- and twelfth-grade "conforming" boys were only children, were more likely to study outside of school, had a higher opinion of school and of teachers, had higher aspirations for more education, had more desire to attend college, had higher-level vocational choices, and attended church frequently.

More tenth-, eleventh-, and twelfth-grade "non-conforming" boys held part-time jobs, owned their own cars, smoked, and stayed around "hangouts." At the tenth- and eleventh-grade levels, these boys had more dates than their "conforming" peers.

Table VI, which is based on the questionnaire study, indicates that "non-conforming" boys generally are over age for their grade level. This, no doubt, is in line with the data which showed them to be poorer students.

Summary

Tenth-, eleventh-, and twelfth-grade "conforming" boys were compared to tenth-, eleventh-, and twelfth-grade "non-conforming" boys in order to determine how the groups differ. The boys in the two groups (matched by IQ) were compared on the characteristics measured by the California Test of Personality, the Thurstone Temperament Schedule, and the California Reading Test. They were compared on the basis of total grades, English grade, physical education grade, and on the number of times they were absent from school, and the number of times they were tardy. Finally, they were asked to respond to a number of questions on a questionnaire designed by the author.

Generally, the "non-conformers" tended to show poorer social adjustment and were more impulsive than the "conforming" boys. The questionnaire results indicated that "conforming" boys come from higher-level socio-economic homes, and have better educated parents. They like school, plan to continue their education, and have higher-level vocational choices. They also attend church frequently. "Non-conformers" held more part-time jobs, owned their own cars, smoked, and stayed around "hangouts".

These findings serve to identify characteristics which differentiate "conforming" from "non-conforming" boys, and thus are of interest to teachers and administrators to aid them in identifying and helping the "non-conformer".

REFERENCES

1. Boynton, P. L., and McGraw, B. H. "The Characteristics of Problem Children," Journal of Juvenile Research, VIII (1943).
2. Coleman, C. T. "The Characteristics of Disciplinary Problem Pupils in High School," School Review, XXXVIII (June 1930).
3. Douglass, Harl R. Organization and Administration of Secondary Schools (New York: Ginn and Company, 1945).
4. Eaton, M. T., et al. "Problem Behavior in School," Journal of Educational Psychology, XLVII (October 1956).
5. Garrett, Henry E. Statistics in Psychology and Education (New York: Longmans Green, 1958).
6. Hales, L. "A Mental Health Survey of One Thousand High School Students in Salt Lake City, Utah," Educational Abstracts, V (July 1940).
7. Krantz, LaVerne. "Administrative Procedure for the Control of Discipline in the High School," American School Board Journal, XCI (September 1935).
8. Lewis, W. D. "Some Characteristics of Children Designated as Mentally Retarded, as Problems and as Geniuses by Teachers," Pedagogical Seminary and Journal of Genetic Psychology, LXX (March 1947).
9. McClure, W. E. "Characteristics of Problem Children Based on Judgments of Teachers," Journal of Juvenile Research, XIII (1929).
10. Mullen, F. A. "Truancy and Classroom Disorder as Symptoms of Personality Problems," Journal of Educational Psychology, XLI (February 1950).
11. Olson, Willard C. Problem Tendencies in Children (University of Minnesota Press: 1930).
12. Owens, Albert A. "The Behavior-Problem

Boy," Journal of Educational Research, XX (October 1929).

13. Snyder, Louise M. "The Problem Child in the Jersey City Elementary Schools," Journal of Education and Sociology, VI (February 1934).
14. Strang, Ruth. "Contributions of Research to Discipline and Control," Thirty-Seventh Yearbook of the National Society for the Study of Education, Part II (1938).
15. Tallent, N. "Behavioral Control in Intellectual Achievement of Secondary School Boys," Journal of Educational Psychology, XVII (December 1956).
16. Tucker, Louise E. "A Study of Problem Pupils," Teachers College Record, XXLI (January 1938).
17. Yourman, Julius. "Children Identified by Their Teachers as Problems," Journal of Educational Sociology, V (1932).

TABLE I

COMPARISONS OF THE DIFFERENCES BETWEEN CONFORMING AND NON-CONFORMING TENTH-, EIEVENTH-, AND TWELFTH-GRADE BOYS ON THE VARIOUS SECTIONS OF THE CALIFORNIA TEST OF PERSONALITY

(Means, Sums of Deviations^2, and "t" values)

	N = 50	Conforming		Non-Conforming		"t"
Grade	Adjustment	Mean	Σd^2	Mean	Σd^2	
10th	Personal	64.4	6182.0	59.1	7256.0	1.848
	Social	60.8	4449.0	51.5	6988.0	3.509*
	Total	125.2	16827.0	110.4	24673.0	2.94*
11th	Personal	65.3	9763.0	63.2	13364.0	.99
	Social	64.8	6710.0	58.9	11125.0	3.17**
	Total	130.2	25071.0	122.1	42924.0	2.19**
12th	Personal	69.8	5775.0	68.2	7572.0	.27
	Social	69.2	4152.0	65.1	6391.0	1.78
	Total	138.3	18131.0	133.3	24093.0	1.05

* .01 level of confidence

** .05 level of confidence

TABLE II

COMPARISONS OF THE DIFFERENCES BETWEEN CONFORMING AND NON-CONFORMING TENTH-, ELEVENTH, AND TWELFTH-GRADE BOYS ON THE VARIOUS SECTIONS OF THE THURSTONE TEMPERAMENT SCHEDULE

(Means, Sums of Deviations^2, and "t" Values)

N = 50		Conforming		Non-Conforming		
Grade	Temperament Trait	Mean	Σd^2	Mean	Σd^2	"t"
10th	Active	9.2	387.04	9.6	468.16	.51
	Vigorous	12.1	460.41	12.7	452.49	.78
	Impulsive	11.0	390.00	11.8	517.84	1.11
	Dominant	6.5	668.25	6.4	665.96	.11
	Stable	8.1	303.21	8.2	307.84	.04
	Sociable	9.1	584.41	9.2	573.64	.15
	Reflective	7.9	463.21	7.0	273.00	1.31
11th	Active	10.2	778.88	11.3	710.08	2.16**
	Vigorous	12.1	880.72	13.9	850.32	3.13**
	Impulsive	11.2	773.08	13.0	686.00	3.25**
	Dominant	8.1	1315.12	7.8	1399.28	.42
	Stable	9.0	841.00	9.0	671.00	.10
	Sociable	10.6	862.91	11.4	599.52	1.53
	Reflective	8.1	995.92	7.5	748.00	1.00
12th	Active	10.2	371.96	10.9	451.24	1.07
	Vigorous	12.2	459.16	13.0	558.00	1.08
	Impulsive	10.3	298.36	12.1	391.24	2.97**
	Dominant	7.6	990.24	8.3	849.65	.69
	Stable	8.6	330.24	7.9	467.24	1.19
	Sociable	10.7	629.96	10.6	560.24	.03
	Reflective	8.7	371.96	7.3	576.56	1.89

** .05 level of confidence

TABLE III

COMPARISONS OF THE DIFFERENCES BETWEEN CONFORMING AND NON-CONFORMING TENTH-, ELEVENTH-, AND TWELFTH-GRADE BOYS ON THE VARIOUS SECTIONS OF THE CALIFORNIA READING TEST

Grade	N = 50 Reading Score	Conforming Mean	Conforming Σd^2	Non-Conforming Mean	Non-Conforming Σd^2	"t"
10th	Vocabulary	50.1	6346.0	40.8	5631.0	3.45*
	Comprehension	32.1	2642.0	28.4	1843.0	2.26*
	Total	82.2	14842.0	69.2	11750.0	3.24*
11th	Vocabulary	52.0	15282.0	50.0	11049.0	.88
	Comprehension	36.7	5052.0	32.8	3970.0	2.13**
	Total	87.5	35477.0	82.8	23919.0	1.37
12th	Vocabulary	62.3	4868.0	57.8	6789.0	1.82
	Comprehension	39.9	1526.0	36.6	2876.0	2.05**
	Total	102.2	10924.0	94.4	15860.0	2.08**

* .01 level of confidence

** .05 level of confidence

TABLE IV

COMPARISONS OF THE DIFFERENCES BETWEEN CONFORMING AND NON CONFORMING TENTH-, ELEVENTH-, AND TWELFTH-GRADE BOYS ON TOTAL GRADE AVERAGE, ENGLISH GRADE, AND PHYSICAL EDUCATION GRADE

Grade	N = 50	Conforming Mean	Conforming Σd^2	Non-Conforming Mean	Non-Conforming Σd^2	"t"
10th***	Total Grade	2.21	14.53	1.51	9.84	5.80*
	English Grade	1.95	28.00	1.34	27.29	3.32*
	Phys. Ed. Grade	2.49	32.25	1.63	37.56	4.15*
11th***	Total Grade	81.0%	1985.0	72.8%	1730.0	9.56*
	English Grade	77.8	3754.0	69.7	3408.0	6.80*
	Phys. Ed. Grade	85.6	1045.0	78.8	3753.0	6.95*
12th***	Total Grade	82.0%	908.0	76.1%	828.0	6.17*
	English Grade	80.0	1755.0	71.9	1484.0	5.89*
	Phys. Ed. Grade	88.0	609.0	81.0	1801.0	5.77*

* .01 level of confidence

** Score based on A =4, B=3, C=2, D=1

***Score based on percentages

TABLE V

COMPARISONS OF THE DIFFERENCES BETWEEN CONFORMING AND NON-CONFORMING TENTH-, ELEVENTH-, AND TWELFTH-GRADE BOYS ON THE NUMBER OF DAYS ABSENT AND THE NUMBER OF TIMES TARDY

Grade	N = 50	Conforming		Non-Conforming		
		Mean	Σd^2	Mean	Σd^2	"t"
10th	Days Absent	4.3	648.89	7.8	1151.84	3.35*
	Days Tardy	2.3	1724.49	2.2	550.49	.15
11th	Days Absent	3.8	1047.68	8.8	2631.28	5.89*
	Days Tardy	2.0	703.00	6.8	1858.88	6.83*
12th	Days Absent	5.5	2479.00	7.6	2256.44	1.32
	Days Tardy	3.5	1867.00	9.4	4244.64	3.30*

* .01 level of confidence

TABLE VI

CHRONOLOGICAL AGE RANGES OF THE TENTH-, ELEVENTH-, AND TWELFTH-GRADE BOYS STUDIED (PERCENTAGES)

Age	Conforming			Non-Conforming		
	10th	11th	12th	10th	11th	12th
15-0 to 15-5	27%	0%	0%	10%	0%	0%
15-6 to 15-11	44	1	0	22	0	0
16-0 to 16-5	15	25	0	27	22	0
16-6 to 16-11	10	39	0	29	22	2
17-0 to 17-5	5	15	23	7	17	11
17-6 to 17-11	0	10	43	5	22	52
18-0 to 18-5	0	4	25	0	7	14
18-6 to 18-11	0	4	9	0	4	7
19-0 to 19-5	0	0	0	0	4	9
19-6 to 19-11	0	0	0	0	0	0
20-0 to 20-8	0	1	0	0	0	5

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 4, December 1962)

Comparisons Between Second-, Third-, and Fourth-Grade Children In the Audio-Lingual Learning of Japanese As a Second Language

ROBERT E. GRINDER
University of Wisconsin

AIKO OTOMO, WINIFRED TOYOTA
University of Hawaii

IN RECENT YEARS attention has focused upon foreign language instruction in the elementary school (FLES) as part of the general effort of educators to produce a more knowledgeable citizenry. Obviously, a person's proficiency in a foreign language assures greater likelihood of his understanding and appreciating societies using the language. But when should an individual start his FL training? Although educators generally agree that children can learn a native-like pronunciation of a foreign language more easily than adolescents or adults (1, 4, 5), research has yet to indicate the best age at which to begin FL instruction (2). Kirch (5), for example, states that the pronunciation of German by first graders is better than that by sixth graders; he calls for FL teaching as early as the first grade. Dunkel and Pillet (4) compared third- and fourth-grade classes, and found nonsignificant differences between them; nonetheless, they recommend starting FL instruction in the third grade because its early introduction would provide more time for teaching. In a nationwide survey of thirty-four school districts where foreign languages are taught in the elementary schools, Levenson (6) found that the frequency distribution of districts, by grade from kindergarten to grade six, at which FLES instruction began was kindergarten, 5; first grade, 7; second grade, 3; third grade, 9; fourth grade, 5; fifth grade, 4; sixth grade, 1. Levenson's (6) data shows that 70% of the districts sampled began FLES instruction by the third grade. Conant (3), however, while recognizing that there may be advantages to FLES instruction, does not "recommend commencing the study of a foreign language prior to the grade seven unless the community demands it, sufficient funds are available, bilingual teachers can be found, and a sequential program can be assured." Administrative problems observed by Conant doubtless are compounded by early training; thus, practically speaking, it seems important that empirical studies be made to determine the best age at which to start FLES instruction.

The present study, accordingly, compared the FL learning of Japanese vocabulary, comprehension, and speech characteristics between second-, third-, and fourth-grade children. The same audio-lingual curriculum was used for one year at each grade; the purpose of the study was to observe whether or not there were any differences between the grades in the learning of Japanese as a second language.

Procedure and Evaluation

FL instruction in Japanese started at the University of Hawaii Elementary School at the beginning of the school year 1959-1960. Questionnaires sent by PTA groups to parents, opinions of FL experts, and community interests were evaluated for three years prior to the adoption of FL teaching in the school curricula. For two compelling reasons, associated with the fairly great cultural influence of Japan in Hawaii, Japanese was chosen as the foreign language: a) parents were eager to have their children learn Japanese, and b) qualified bilingual teachers of Japanese were regular members of the elementary school faculty.

Instruction in Japanese began in the second-,

---

* Footnotes will be found at the end of the article.

third-, and fourth-grade classes in October and concluded late in May. Eighty-six sessions were held. Except for the lapse of a month between semesters, one fifteen-minute session was held every school day. All of them were conducted in the morning in the classrooms. Characteristics of the subjects by age, grade, and sex are given in Table I. Subjects were unselected in terms of intelligence and social class. Two bilingual teachers were used; each teacher conducted a class at each grade. Therefore, each teacher taught three classes every morning. Both teachers had twelve years of training at a Japanese language school, and both were experienced elementary school teachers. At the time of the study, they were supervisors of practice teachers.

The FL instruction in Japanese was based upon an audio-lingual approach. Subjects were encouraged to listen and to imitate the teacher's speech. Prior to beginning instruction, a curriculum guide was developed by the two bilingual teachers. The guide stressed the basic vocabulary that the subjects would need for conversation about their families, classroom objects, colors, holidays, etc. The learning of Japanese was stimulated by several techniques. An effort was made to include in equal proportion: direct conversation, songs, story-telling, dramatizations, dancing, games, drawings on the blackboard, flannelboard cutouts, pictures, phonograph records, and tape recordings. Insofar as possible, every class received identical treatment. English was not used during the fifteen-minute FL session. In order that the subjects could become accustomed to a voice other than that of their teacher, occasionally they were asked to answer the questions of an unknown Japanese man, whose voice they heard from a tape.

Subjects were evaluated at the end of the school year in terms of vocabulary, comprehension and speech characteristics. Measures of subjects' vocabulary and comprehension were obtained with paper-and-pencil tests; measures of subjects' speech characteristics were obtained by rating from tapes their previously recorded verbal behavior.

In constructing the vocabulary and comprehension tests, the two bilingual teachers first listed the approximately 200 words that had been used in the curriculum. Then the words were divided, by estimate, into quarters on the basis of the frequency of their usage. From each of the quarters, ten words were selected; five were assigned to the vocabulary test and five to the comprehension test. Each test, therefore, was comprised of twenty items. At each stage of test construction, words were randomly distributed. The tests were administered in the subjects' classrooms on separate days by the FL teachers. English was used to communicate instructions for taking the tests. Subjects were unaware that they were being tested. A subject's score for each test was the number of items that he correctly answered.

Subjects responded to both tests by choosing from four pictures the one that best represented the test item, and by marking it with an "X". The danger of subjects being confused by the pictures upon confrontation was minimized by printing the four choices for each item on a narrow strip of paper. The strips were stapled one on top of the other so that a subject saw no more than one four-choice set at a time. In assessing subjects' vocabulary, the teacher read in Japanese the list of twenty words, one at a time. After each word, subjects marked an "X" over the picture that best described the word. In assessing subjects' comprehension, the teacher read in Japanese brief descriptions of simple situations; following each description a short question was asked about it. After each description-question, subjects marked an "X" over the picture that depicted the activity which best answered the question.

Subjects' articulation of Japanese was assessed in terms of the degree to which they accurately recited ten short sentences. The sentences were constructed to elicit particularly nine basic characteristics of Japanese speech. The characteristics constituted those important for effective articulation of Japanese. Those that were assessed are: pause; accent; silent vowels; long vowels; dental consonants, d, t, n; ts; and "flapped" r. Each of the nine speech characteristics appeared approximately six times in the ten sentences. The sum of them was 51.

Subjects were sent individually from their classroom to the experimental room, which was a quiet, small room furnished with a table, two chairs, and a room-divider. The tape recorder was hidden behind the room-divider; the microphone laid under some flowers in a basket on the table. During the experimental session, only one experimenter and a subject were in the room. Upon entering the room, the subject was put at ease by the experimenter (who was one of the bilingual teachers), with some familiar Japanese greetings. Instructions were given in English; then, the ten sentences were read one at a time in Japanese. The experimenter repeated every sentence twice. After hearing a sentence for the second time, the subject recited it. The Japanese words and phrases used in the sentences were unfamiliar to the subjects. As with the vocabulary and comprehension tests, the subjects did not know that they were being evaluated.

Accuracy of subjects' recitation of the ten sentences was rated by a bilingual graduate student in Japanese. Subjects were given a separate score for each of the nine speech characteristics. Each score represented the number of times that a given speech characteristic had been correctly

TABLE I

MEANS AND STANDARD DEVIATIONS OF SUBJECTS' AGE IN MONTHS AT END OF SCHOOL YEAR BY SEX AND GRADE

Grade	Males			Females			Total		
	Mean	S.D.	N	Mean	S.D.	N	Mean	S.D.	N
2	96.8	4.05	15	97.0	4.78	14	96.9	4.42	29
3	109.4	4.43	13	107.0	4.38	16	108.1	4.56	29
4	119.0	3.74	13	119.2	3.89	17	119.1	3.83	30

COMPARISONS BETWEEN TWO TEACHERS AND THE AUDIO-LINGUAL LEARNING OF JAPANESE BY ELEMENTARY SCHOOL CHILDREN

Dependent Variable	Teacher 1 (N = 40)	Teacher 2 (N = 48)	x^{2*}	p^{+}
Vocabulary			1.63	<.30
High	18	30		
Low	22	18		
Comprehension			.99	<.50
High	21	19		
Low	19	29		
Pause			.36	<.70
High	19	27		
Low	21	21		
Accent			.13	<.80
High	29	32		
Low	11	16		
Silent vowels			1.17	<.30
High	23	34		
Low	17	14		
Long sounds			.11	<.80
High	24	26		
Low	16	22		
d			.33	<.70
High	25	26		
Low	15	22		
n			1.11	<.30
High	24	35		
Low	16	13		
t			.93	<.50
High	20	18		
Low	20	30		
ts			5.32	<.05
High	11	26		
Low	29	22		
r			6.70	<.01
High	27	18		
Low	13	30		

---

* Corrected for continuity.

+ One degree of freedom, two-tail.

articulated. Accuracy of the rater was checked on several occasions by one of the bilingual teachers; in every instance, the proportion of agreement was above 90%.

Results

Children in the three classes who had attended Japanese language school, whose parents spoke Japanese, or who lived in households where they heard Japanese spoken, were excluded from the sample. Of 148 children, 60 were found to have had previous experience with Japanese; therefore, 88 subjects, initially unsophisticated in Japanese, were used in the statistical analyses. The frequency distributions of all the dependent variables, i.e., the vocabulary test, the comprehension test, and the nine speech characteristics, proved to be skewed and unamenable to successful transformations. For this reason the analyses were made by means of chi-square. Each dependent variable was divided as near the median as the distribution allowed. Subjects were thus separated into low and high groups on each measure.

First, tests were conducted between ethnic background (the sample comprised Caucasian, Chinese, Filipino, Hawaiian, Japanese, and Korean children), and sex, respectively, and each of the dependent variables to determine whether or not these antecedents should be controlled in subsequent analyses. All the tests of association were nonsignificant. Apparently both ethnic background and sex are not important factors in children's learning of Japanese.

Table II shows comparisons between the two bilingual teachers and subjects' performances on, respectively, the tests of vocabulary, comprehension, and speech characteristics. Significant teacher differences appeared with the "flapped" r and the ts sounds. Subjects of Teacher 1 are more effective in articulating the sound r; subjects of Teacher 2 are more effective in articulating the sound ts. None of the other tests of teacher differences was significant.

Table III shows comparisons between the three grades and each of the dependent variables. Higher grade is significantly associated with higher learning of: comprehension (p .001), silent vowels (p .01), t (p .05), and ts (p .05). A trend appears for higher grade to be related to high vocabulary (p .10) and high d (p .20).

Table IV shows the analyses of Table III separated by teacher. Although the power of the chi-square test is reduced by the smaller N, the data of Table IV indicate consistently that "high" subjects of both teachers are likely to be from the fourth grade. Since Table II shows that teacher differences may exist in FLES instruction, an inspection of the cell entries for the tests of association for the sounds r and ts was made. Although Teacher 1 is significantly more proficient than Teacher 2 with sound r, the only interaction of the sound with grade appears with Teacher 2, who was more effective with grade 3 than with either grade 2 or 4. Grade seems to interact with both teachers in the development of the sound ts. Teacher 1, who was shown in Table II to be less effective than Teacher 2 in teaching the sound ts, appears unavailing with grade 2; Teacher 2 seems proficient with grade 4.

Discussion

The evidence presented in Table II reveals that teacher variation in the quality of FLES instruction may yield differences in children's FL learning. Clearly the classes of the two teachers differed in respect to the r and the ts sounds. Perhaps it is significant that these Japanese speech characteristics are among those that are virtually nonexistent in English. Apparently the two bilingual teachers, equally competent otherwise, differed in the degree to which they had mastered these sounds.** Consequently, Teacher 1's greater command of the r sound enabled her to be more efficient in teaching it; similarly, Teacher 2 was more efficient with the sound ts. The data are consistent with the notion that excellence in FLES learning is positively correlated with teacher competence.

The findings reported in Table III show strong, positive relationships between grade and the audio-lingual learning of Japanese comprehension and several speech characteristics. The strength and consistency of the associations contrast with the experience of Kirch (5) who found an inverse relationship between elementary school age and ability to reproduce foreign sounds. However, Kirch's (5) finding and his assertion that FLES instruction should begin in the first grade, rather than the sixth grade, are based upon results obtained under conditions that were likely to have facilitated the superior performance of his first graders. Some children in the first grade were motivated to learn German because they wanted to communicate with their German-speaking parents and grandparents; doubtless the children had learned some pronunciation and accent skills at home. Furthermore, the first-grade FL curriculum, which stressed concepts about the immediate social environment, employed only conversational methods; on the other hand, the sixth-grade FL curriculum, which included the geography of Germany and Austria, plus a drama written in German by the students, employed both conversational and reading methods. Thus, the first grade, because of the higher motivation, the situational relevance of the curriculum, and the emphasis upon conversation, probably had several advantages over the sixth grade in learning German pronunciation. The results of the present

TABLE III

COMPARISONS BETWEEN ELEMENTARY SCHOOL CHILDREN IN THE SECOND, THIRD, AND FOURTH GRADES AND THE AUDIO-LINGUAL LEARNING OF JAPANESE

Dependent Variable		Grade		x^2	p^+
	2 (N = 29)	3 (N = 29)	4 (N = 30)		
Vocabulary				5.01	<.10
High	12	15	21		
Low	17	14	9		
Comprehension				16.77	<.001
High	6	12	22		
Low	23	17	8		
Pause				1.82	<.50
High	18	13	15		
Low	11	16	15		
Accent				.77	<.70
High	22	19	22		
Low	7	10	10		
Silent vowels				11.45	<.01
High	13	18	26		
Low	16	11	4		
Long sounds				1.31	<.70
High	18	14	18		
Low	11	15	12		
d				3.84	<.20
High	17	13	21		
Low	12	16	9		
n				1.43	<.50
High	17	18	22		
Low	12	9	8		
t				7.60	<.05
High	10	9	19		
Low	19	20	11		
ts				6.54	<.05
High	7	13	17		
Low	22	16	13		
r				.78	<.50
High	14	17	14		
Low	15	12	16		

[+] Two degrees of freedom, two-tail.

TABLE IV

COMPARISONS BETWEEN ELEMENTARY SCHOOL CHILDREN IN THE SECOND, THIRD, AND FOURTH GRADES SEPARATED BY TEACHER AND THE AUDIO-LINGUAL LEARNING OF JAPANESE

	Teacher 1 Grade					Teacher 2 Grade				
Dependent Variable	2 (N=14)	3 (N=14)	4 (N=12)	x^2	p^+	2 (N=15)	3 (N=15)	4 (N=18)	x^2	p^+
Vocabulary										
High	3	8	7	4.80	<.10	9	7	14	2.78	<.30
Low	11	6	5			6	8	4		
Comprehension										
High	4	8	9	5.72	<.10	2	4	13	13.34	<.01
Low	10	6	3			13	11	5		
Pause										
High	9	7	3	4.00	<.20	9	6	12	2.45	<.30
Low	5	7	9			6	9	6		
Accent										
High	12	10	7	2.44	<.30	10	9	13	.48	<.80
Low	2	4	5			5	6	5		
Silent vowels										
High	6	8	9	2.72	<.30	7	10	17	9.22	<.01
Low	8	6	3			8	5	1		
Long sounds										
High	10	7	7	1.36	<.70	8	7	11	.67	<.80
Low	4	7	5			7	8	7		
d										
High	7	9	9	1.72	<.50	10	4	12	6.62	<.05
Low	7	5	3			5	11	6		
n										
High	7	9	8	.88	<.70	10	11	14	.48	<.80
Low	7	5	4			5	4	4		
t										
High	5	6	9	4.40	<.20	5	3	10	4.56	<.20
Low	9	8	3			10	12	8		
ts										
High	1	6	4	4.72	<.10	6	7	13	3.88	<.20
Low	13	8	8			9	8	5		
r										
High	10	8	9	1.04	<.70	4	9	5	4.66	<.10
Low	4	6	3			11	6	13		

+ Two degrees of freedom, two-tail.

study, obtained under controlled conditions with a standardized curriculum, clearly indicate that FLES instruction in Japanese is more effective in the fourth grade than in the second.

The data of Table IV also support the inference that children in the fourth grade are more effective than those in the second grade in the audio-lingual learning of Japanese. The statistical relationships that resulted are roughly the same for both teachers, and there is no evidence to suggest that one teacher's effort contributed much more than the other's to the results. The observation drawn from Table IV, albeit inconclusive, that teacher differences may interact with grade suggest the possibility that poor FL teachers may be especially ineffectual with the lower grades, whereas good FL teachers may be particularly effective with higher grades.

In spite of the evidence of this study that suggests that older elementary-age children profit more from FLES instruction, there is the question as to what FL performance might be in the fourth grade if the children could be continuously given FL training from the second grade. Presumably there would be a positive difference between those who had three years' training and those who had just one year, but the difference that might occur from the additional years of training must be justified in terms of teacher time, curriculum, and administrative expense. Since the findings of this study indicate that older children's FL performance, relative to the amount of time given to instruction, is more accelerated than that of younger children, any difference that accrues from early training will have to be fairly extensive before it is practical.

On the basis of the findings reported here, this study suggests that FLES instruction should be introduced later rather than earlier in the elementary school curriculum. The suggestion is contrary to current practice in the United States (6), but it is in accord with the recommendation of Conant (3) that FL instruction should begin in the seventh grade.

Summary

Children in the second, third, and fourth grades were taught the same audio-lingual curriculum in Japanese for one school year. The purpose of the study was to investigate differences between the grades in the learning of Japanese as a second language. Two bilingual teachers conducted the FL sessions; each teacher was assigned to one class at each grade. Subjects were evaluated in terms of vocabulary, comprehension, and speech characteristics. Neither ethnic background nor sex was found to be related to FL learning. However, empirical tests showed that the two teachers differed significantly in teaching two Japanese speech characteristics, and that the fourth-grade subjects were consistently and significantly more likely than those in the lower grades to be high in comprehension and several speech characteristics. When the sample was separated by teacher, results were generally the same: fourth-grade subjects taught by either teacher were more effective than those in the second or third grades in learning Japanese as a second language.

FOOTNOTES

* The authors wish to thank Miss Gladys M. Traut, Principal of the University Elementary School, University of Hawaii, for her cooperation in making this research possible, and Herbert B. Weaver, for his helpful criticism of the manuscript. The authors gratefully acknowledge the technical assistance of Norito Fujioka, who carefully checked the curriculum for the accuracy of direct and idiomatic translations from English to Japanese.

** An informal test of the teachers' pronunciation of the r and ts sounds indicated that variations existed: Teacher 1 appeared to enunciate r better than Teacher 2, and Teacher 2 appeared to enunciate ts better than Teacher 1.

REFERENCES

1. Carroll, J. B. "Foreign Languages for Children — What Research Says," The National Elementary Principal, XXXIX (1960), pp. 12-15.
2. Carroll, J. B. "Wanted: A Research Basis for Educational Policy on Foreign Language Teaching," Harvard Educational Review, XXX (1960), pp. 128-140.
3. Conant, J. B. Education in the Junior High-School Years (Princeton, New Jersey: Educational Testing Service), 1960.
4. Dunkel, H. B., and Pillet, R. A. "A Second Year of French in Elementary School," The Elementary School Journal, LVIII (1958), pp. 143-151.
5. Kirch, M. S. "At What Age Elementary School Language Teaching?" The Modern Language Journal, XXXX (1956), pp. 399-400.
6. Levenson, S. "The Current Status of FLES Teaching in the U. S.," Phi Delta Kappan, XXXXII (1961), pp. 265-268.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 4, December 1962)

A Study of the Effectiveness of Certain Techniques of Reading Instruction in Developing the Ability of Junior High-School Students to Conceptualize Social Studies Content

MARY WITT
Florida State University

THE CONCEPT APPROACH to the teaching of social studies is one way to help students develop critical thinking. The meaning of a social-studies concept, as used in this study, is an idea which contains many levels of depth and breadth, and which may be applied to the past, present, and future, for example, "World Interdependence."

An exploratory study designed to identify techniques and to test the validity and reliability of certain evaluative instruments was conducted in the summer of 1959 with a small group of superior students. As a result of this work, a body of techniques was developed which appeared to be promising for classroom instruction in reading and in developing concepts in the social studies.

The purpose of this study is to determine the amount of observable change which was brought about in reading, in social studies in general, and in the developing of ten specific social studies concepts through classroom instruction and guidance.

Procedure

The 62 seventh-grade students of the University School were divided into two equated groups. Grouping was arrived at on the basis of the average of the students' scores on the California Language and Social Studies Tests, Intermediate Form CC. These tests were administered to this group of students in May 1959 near the end of the sixth grade. The resulting scores indicated a range from grade 3.9 to grade 9.8 (Table I). Since this is a report of parallel studies, it was not necessary for the subjects to be precisely matched, though the means were in fact almost identical. Twenty-nine subjects were used on all data. Two additional students in each group were dropped due to absence at the time the first Concept Test was administered; also, the experimental work on concept development was well under way before they returned to school.

As a part of the school testing program, the following tests were administered during the first month of school:

1. California Reading Test, Intermediate Form BB
2. California Language Test, Intermediate Form BB
3. California Short-Form Test of Mental Maturity, Junior High-School Level, 1957 S Form

The analysis of the scores from these tests indicated a range from grade 3.7 to grade 11.5 (Table I).

For the purpose of this study the following tests were administered in October:

1. STEP Social Studies Test, Form 3A
2. Iowa Silent Reading Test, Form AM (Revised)

In order that test scores might reveal the improvement of the students on the higher levels of reading, the Advanced Test was administered to those whose scores on the California Reading Test were 7.5 or above. The Elementary Test was used with those whose scores were below 7.5. Sixteen students in each group took the Advanced

TABLE I

FREQUENCY DISTRIBUTION OF CALIFORNIA ACHIEVEMENT TEST SCORES

Class Intervals (In months)		Average Language and Social Studies, May 1959		Reading, September 1959	
		Group I	Group II	Group I	Group II
115 to	119	0	0	1	0
110	114	0	0	0	0
105	109	0	0	0	0
100	104	0	0	2	0
95	99	1	1	1	5
90	94	1	1	4	4
85	89	2	2	1	2
80	84	3	2	5	2
75	79	4	6	2	3
70	74	4	1	3	0
65	69	7	6	2	2
60	64	1	2	2	3
55	59	0	2	0	2
50	54	4	3	4	2
45	49	1	1	1	2
40	44	0	0	0	2
35	39	0	2	1	0
30	34	1	0	0	0
		N=29	N=29	N=29	N=29
		Range 71 Median 71	Range 66 Median 69	Range 91 Median 76	Range 71 Median 76

TABLE II

IOWA SILENT READING TEST, OCTOBER - FEBRUARY

	Group I		Group II	
	October	February	October	February
M	154.69	158.66	151.86	154.97
σ	15.81	18.5	17.06	16.4
σ_M	2.94	3.34	3.17	3.05
$D_{M_2 - M_1}$		3.97		3.11
r_{12}		.91		.96
SE_D		1.38		.90
t		2.88		3.44

N = 29

Where $t \geq 2.47$, $P \leq .01$

TABLE III

STEP SOCIAL STUDIES TEST, OCTOBER - FEBRUARY

	Group I		Group II	
	October	February	October	February
M	263.1	267.41	260.48	264.66
σ	15.44	15.85	12.24	15.39
σ_M	2.87	2.92	2.27	2.85
$D_{M_2-M_1}$	4.31			3.15
r_{12}	.89			.95
SE_D	1.36			.99
t	3.17			3.08

N = 29

Where $t \geq 2.47$, $P \leq .01$

TABLE IV

IOWA SILENT READING TEST, FEBRUARY - JUNE

	Group I		Group II	
	February	June	February	June
M	158.66	163.34	154.97	158.79
σ	18.5	20.46	16.4	15.27
σ_M	3.34	3.79	3.05	2.84
$D\ M_2-M_1$		4.86		3.82
r_{12}		.94		.95
SE_D		1.31		.95
t		3.57		4.25

N = 29

Where $t \geq 2.47$, $P \leq .01$

Test; thirteen students in each group took the Elementary Test.

At the end of the first semester, before beginning the experimental work on "Concept Teaching," the following tests were administered:

1. STEP Social Studies Test, Form 3B
2. Iowa Silent Reading Test
    a. Elementary Test, Form BM (Revised)
    b. Advanced Test, Form BM (Revised)
3. Teacher-Made Concept Test*

At the end of the second semester the following tests were administered:

1. STEP Social Studies Test, Form 3A
2. Iowa Silent Reading Test
    a. Elementary Test, Form CM
    b. Advanced Test, Form CM
3. Teacher-Made Concept Test

The classes with which this study was conducted met during a two-hour block of time scheduled for the Language Arts and Social Studies in the Junior High School. Two skills were particularly stressed through the organized reading instruction and, in turn, these skills were applied to the development of the social studies program. Stress was placed upon the ability 1)to organize through outlining and 2)to draw conclusions from factual materials. Many skills and understandings are involved in the further development of these specific techniques.

Although other units were developed during the school year, this report is concerned with units on Latin America and Canada and their relationship to the United States and to the World. During the first semester, Group I studied Latin America while Group II studied Canada. In both groups, the usual unit organization and methods were employed.

At the beginning of the second semester, the bodies of content were reversed with the two groups; i. e., Group I studied Canada while Group II studied Latin America. The method employed for this period was different. It was consistently focused upon the development of the following ten selected concepts: 1) health, 2) world interdependence, 3) natural resources, 4) education, 5) moral and spiritual values, 6) self-government, 7) social responsibility, 8) culture as revealed in creative expression, 9) technological advances, and 10) the spirit of pioneering. As mentioned above, the Concept Test was administered at the beginning and at the end of the second semester.

Findings

Table II is a tabulation of the Iowa Silent Reading Test administered to Group I and to Group II in October and in February.

For Group I, the difference between the October mean and the February mean was 3. 97. For Group II, the difference between the October mean and the February mean was 3. 11. The t score of each group indicates that a significant gain was made in reading during this period of time by each group of students.

Table III is a tabulation of the STEP Social Studies Test administered to Group I and to Group II in October and in February.

For Group I, the difference between the mean of the October scores and the mean of the February scores was 4. 31. For Group II, the difference between the mean of the October scores and the mean of the February scores was 3. 15. As indicated in Table IV, a significant gain was made in social studies by each group of students during this period of time.

Table IV is a tabulation of the Iowa Silent Reading Test administered to Group I and to Group II in February and in June.

For Group I, the difference between the mean of the February scores and the mean of the June scores was 4. 86. For Group II, the difference between the mean of the February scores and the mean of the June scores was 3. 82. As indicated in Table IV, a significant gain was made in reading by each group during the second semester of the school year.

Table V is a tabulation of the STEP Social Studies Test administered to Group I and to Group II in February and in June.

For Group I, the difference between the mean of the February scores and the mean of the June scores was seven. For Group II, the difference between the mean of the February scores and the mean of the June scores was 7. 23. The t scores, as shown in Table V, indicate that a significant gain was made in social studies by each group during the second semester of the school year.

In comparing the gains made by Group I and by Group II between the October to February period and the February to June period, it was found that a greater gain was made in reading by each group in the second period than in the first period. However, this gain or difference in gain was not statistically significant. In the social studies, a greater gain was made in the period from February to June than from October to February, which gain was statistically significant.

---

* This test was prepared by the teacher with the assistance of Dr. Willard Nelson, Psychology Department, Florida State University, and Mr. Herbert Hughes, doctoral intern in the School of Psychology.

TABLE V

STEP SOCIAL STUDIES TEST, FEBRUARY - JUNE

	Group I		Group II	
	February	June	February	June
M	267.41	274.41	264.66	271.89
σ	15.85	15.40	15.39	15.87
σ_M	2.92	2.87	2.85	2.95
$D_{M_2-M_1}$		7.		7.23
r_{12}		.89		.89
SE_D		1.36		1.36
t		5.31		5.15

N = 29

Where $t \geqq 2.47$, $P \leqq .01$

TABLE VI

TEACHER-MADE CONCEPT TEST, FEBRUARY - JUNE

	Group I		Group II	
	February	June	February	June
M	4.76	7.14	3.86	5.76
σ	2.37	1.89	1.73	1.97
σ_M	.44	.35	.32	.37
$D_{M_2-M_1}$		2.38		1.90
r_{12}		.97		.79
SE_D		.34		.23
t		7.00		8.33

N = 29

Where $t \geqq 2.47$, $P \leqq .01$

TABLE VII

SIGN TEST APPLIED TO TEACHER-MADE CONCEPT TEST, FEBRUARY - JUNE

	Group I Rating on Insight into Social Studies Concepts			Group II Rating on Insight into Social Studies Concepts		
Students	June	February		June	February	
1.	49	39	+ 10	35	30	+ 5
2.	30	30	0	38	35	+ 3
3.	33	32	+ 1	32	30	+ 2
4.	35	30	+ 5	39	30	+ 9
5.	35	32	+ 3	33	30	+ 3
6.	30	31	- 1	34	33	+ 1
7.	37	32	+ 5	42	32	+ 10
8.	30	32	- 2	33	32	+ 1
9.	30	30	0	34	32	+ 2
10.	41	32	+ 9	41	35	+ 7
11.	41	40	+ 1	51	42	+ 9
12.	32	30	+ 2	39	35	+ 4
13.	33	35	- 2	33	30	+ 3
14.	43	38	+ 5	54	52	+ 2
15.	39	36	+ 3	37	34	+ 3
16.	32	35	- 3	32	30	+ 2
17.	49	32	+ 17	37	32	+ 5
18.	47	41	+ 6	39	34	+ 5
19.	39	34	+ 5	35	33	+ 2
20.	46	39	+ 7	36	38	- 2
21.	76	44	+ 32	32	36	- 4
22.	40	30	+ 10	41	41	0
23.	40	32	+ 8	38	35	+ 3
24.	41	38	+ 3	40	40	0
25.	39	33	+ 6	31	31	0
26.	40	32	+ 8	46	38	+ 8
27.	36	36	0	33	34	- 1
28.	37	33	+ 4	43	39	+ 4
29.	42	36	+ 6	41	34	+ 7
	N = 26	Z = 3.4		N = 26	Z = 3.8	
	When $Z \geq 3.4$ or $Z \leq -3.4$, P = .00			When $Z \geq 3.8$ or $Z \leq -3.8$, P = .00		

Table VI is a tabulation of the Teacher-Made Concept Test administered to Group I and to Group II in February and in June.

For Group I, the difference between the February mean and the June mean was 2.38. For Group II, the difference between the February mean and the June mean was 1.9. The t scores, as shown in Table VI, indicate that significant gains were made by both groups of students in the development of social studies concepts.

Treatment of the Subjective Data Included in the Teacher-Made Concept Test

Each of the four statements listed under the descriptive material of the test is true, but each represents a different level of abstraction. In evaluating the written comments of the students, three aspects of conceptualization were considered: 1) generalization, 2) discrimination, and 3) extension. Generalization ranges from merely restating facts to the formulation of a principle. Discrimination ranges from non-selective application to a selective application to the chosen answer. Extension extends from simply restating or repeating the idea of the setting presented in the printed material to other relevant applications of the idea implied in the selection.

Each test was read to determine the extent of the three aspects of conceptualization, using the following scale:

1. Did nothing, mixed up
2. Made some effort
3. Average amount of response
4. More than average

Each item was scored on this basis and the results tabulated in Table VII.

Table VII is a tabulation of the Sign Test as applied to the subjective data included in the reasons given by the students in Group I and in Group II for their selection of the answers chosen in the Teacher-Made Concept Test.

Since a quantitative measure was not feasible to evaluate the subjective data of the Teacher-Made Test, it is possible to establish whether or not the conditions before and after the treatment were different by the application of the Sign Test. As shown in Table VII, twenty-two students in Group I and twenty-three students in Group II improved in their ability to relate their comments to the concepts stressed in social studies between February and June.

Since there are no data on concept formation for the October to February period, it is not assumed that the subjects made no progress in the development of concepts in social studies during that time. However, the results of both the objective and the subjective phases of the Teacher-Made Concept Test indicate that significant gains were made by both groups of students in their ability to conceptualize social studies content.

Conclusions

1. Significant gains were made in both reading and in social studies during each semester.
2. A significantly greater gain was made during the second semester period than during the first semester period in social studies.
3. Teaching by the method of ten selected concepts produced a statistically significant gain in concept scores for both groups of students.

Implications

1. The concept approach to the teaching of social studies is a desirable method by which critical thinking may be developed.
2. Specific skills and techniques which are stressed in reading are effective when applied to the social-studies program. However, increased skill in reading does not necessarily mean increased skill in the conceptualization of social-studies content.
3. The concept approach, as used in this study, provides for a wide range of differences in interests, in abilities, and in levels of abstraction represented in a seventh-grade group of students.

REFERENCES

1. Anderson, Howard R. (ed.) Teaching Critical Thinking in the Social Studies, Thirteenth Yearbook of the National Council for the Social Studies (1942).
2. Dressel, Paul I., and Mayhew, Lewis B. Critical Thinking in Social Science (Dubuque, Iowa: Wm. C. Brown Co., 1954).
3. Hyram, George H. "An Experiment in Developing Critical Thinking in Children," Journal of Experimental Education, XXVI (December, 1957), pp. 125-132.
4. McCutchen, Sam (ch.) A Guide to Content in the Social Studies, Report of NCSS Committee on Concepts and Values, National Council for the Social Studies (1958).
5. Spache, George D., and Berg, C. The Art of Efficient Reading (The MacMillan Co., 1955).
6. Stewart, L. Jane, Heller, Frieda M., and Alberty, Elsie J. Improving Reading in the Junior High School (New York: Appleton-Century-Crofts, Inc., 1957).

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 4, December 1962)

Readability Versus Reading Ability

WILBUR R. MILLER
Purdue University

THE INCREASED complexity of our society has resulted in a trend toward increasing the amount and changing the type of information available for inclusion as content in many subject matter fields. The industrial and mechanical phase of our society, represented by industrial arts in the secondary school, is certainly no exception to this trend.

Those responsible for choosing content in industrial arts are placing more and more emphasis upon the acquisition of technical and related information as developments occur in the form of new processes, materials, tools, and machines which increase the level of technology.

There are various means of presenting technical and related information, but perhaps the one most commonly used is the textbook. The importance of reading to the acquisition of technical and related information presented by the textbook is obvious, but should not be minimized. In fact, the ability to read and comprehend instructional materials prepared for a given grade level or subject matter area has long been recognized as one of the important criteria for school success.

Students who read poorly find the road to school success extremely difficult. The possible frustrations brought about by reading difficulties often result in failure and rejection of school. This failure and rejection often cause students to drop out of school and in turn may lead to delinquency and other major school and community problems.

The causes of reading difficulties are many and varied, and the solutions to reading problems are not simple. One of these problems, that of incompatibility between the readability of instructional materials and the reading abilities of students was the focal point of the investigation*[1] reported here.

The Research Project

This research study attempted to find the degree of compatibility between the readability of general shop textbooks in industrial arts and the reading abilities of ninth grade industrial arts students. Answers to the following three questions were sought:

(1) What are the rated readability[2] levels of selected general shop textbooks?

(2) What are the reading abilities of a selected group of ninth grade industrial arts students?

(3) What is the relationship between the reading abilities of a selected group of ninth grade students and the readability of certain general shop textbooks used in the junior high school industrial arts program?

Related Research

Since 1940, there has been a significant increase in the number of research studies using readability formulae to assess the reading difficulty of school textbooks and references.

In an investigation by Porch[3], the Lorge formula of readability was applied to social studies textbooks used in grades three to twelve. The findings indicated that although a progression of difficulty existed from one grade to another, the progression was not consistent. This progression could not provide for the normal growth in reading ability of the students. The greatest departure from uniform progression was between the sixth and seventh grade textbooks, and between the ninth and tenth grade textbooks.

Several studies of this type have been done in science, English, social studies and literature; however, examination of the professional literature indicates a general neglect of certain subject areas with regard to the problems involving reading difficulty of instructional material.

No doubt an awareness of the lack of research in these areas prompted Siceluff[4], in his study of vocabulary loads, to point up the fact that little is known about the vocabulary difficulties involved in such subjects as home economics, agriculture, art, speech, industrial arts, and dramatics. Siceluff suggested that an understanding of reading difficulties involved in these various subjects would be of value in directing pupils into elective courses in which they could achieve a higher degree of reading success.

---

*All footnotes will be found at end of article

Tubbs used the Dale-Chall and Flesch formulae to study the reading difficulty of thirty geography textbooks. The results obtained through the application of these formulae correlated positively with the judgment of teachers, who had used the textbooks. In addition, a correlation of .80 existed between the results obtained through application of each formula. There was a mean difference of one-half year between the grade placements.[5]

Anderson[6] studied a selected group of general business textbooks to ascertain readability levels. This study employed the Yoakum and Flesch formulae, but did not attempt to compare the textbook readability ratings with the measured reading abilities of secondary students.

One of the most recent studies was made by Wyatt[7]. This study was made to ascertain the readability of selected textbooks for elementary social studies and to compare this readability with the measured reading abilities of selected fifth grade pupils. In terms of the factors studied and the procedures used, the study by Wyatt was more closely related to the research project being reported here than any other study reviewed by the writer.

Procedure

Since this study dealt with the reading difficulty of textbooks within a specific area, it was necessary to set up a criterion for the selection of textbooks to be analyzed. Due to the fact that the study was limited to the general shop area of industrial arts, only textbooks that were organized around three or more industrial arts activities (e.g., drawing, wood, metal, electricity, plastics, leather, etc.) were selected for use in the research project. All of these had been written or revised within the past ten years.

In addition to the selection of textbooks, it was necessary to choose a method of assessing the readability level of each of the textbooks. After a general investigation of readability formulae, the writer selected the Dale-Chall and the Flesch formulae for use in finding the rated readability of general shop textbooks.

In order to compare the readability of general shop textbooks with the reading abilities of students, it was necessary to select and test a group of general shop students. Ninth grade students were selected for this phase of the study since the ninth grade is the top level of a three year sequence in which general shop instruction is most often given. These students attended schools located within a twenty-three county area of central Missouri which had ten or more students enrolled in ninth grade general shop.

The silent reading comprehension section of the Iowa Tests of Basic Skills[8] and the Lorge-Thorndike Test of Intelligence[9] were administered to the 411 students involved in the study.

After these data had been obtained, a statistical analysis was made to compare the rated readability levels of the textbooks with the measured reading abilities of the students.

Textbook Readability

An application of the readability formulae to the five textbooks revealed a wide range of readability levels within the sections of each book as well as between the sections.[10] In fact, this range was found to be as great as eleven grade levels in one textbook (fifth grade to college graduate) and no smaller than seven grade levels (fifth grade to twelfth grade) in the textbook which had the smallest range of variability. This information is often unreported when the average of samples from a textbook are revealed.

As indicated by Table I, there was close agreement between the formulae with regard to the reading difficulty of the selected textbooks. A comparison of the readability ratings of the five textbooks indicated a total range of only two grade levels between Book A, which was rated as having the highest level of reading difficulty (tenth grade), and Book B, which was rated as having the lowest level of reading difficulty (eighth grade). The difference between results obtained through the application of the two readability formulae was small. The greatest deviation between formula ratings, one and one-half grade levels, occurred in Book E. The other textbooks, Book C and Book D, were rated slightly more difficult to read by the Dale-Chall formula than the Flesch formula.

Further analysis of the data gathered through application of the readability formulae revealed that sections of the general shop textbooks dealing with electricity, plastics, and home maintenance were rated as more difficult to read than the drawing and planning, metal working, and woodworking sections. The difference comes about primarily through the vocabulary included in these areas. This may imply a need for word study and vocabulary building which should preceed and accompany student use of a textbook in these areas.

Measurement of Reading Abilities

The readability ratings of the textbook or instructional material may be of only limited significance unless the reading abilities of the various individuals, who are to use the material, are known. This fact prompted the investigator to select a group of general shop students, who might be required to use one of these five general shop books as a basic text in industrial arts, and ascertain the reading ability level of the group.

The testing was done during the latter part of the school year; therefore, a student who was reading

TABLE I

READABILITY, IN GRADE LEVELS, OF TEXTBOOKS AS RATED BY DALE-CHALL AND FLESCH FORMULAS

Book Code	Dale-Chall	Flesch	Difference	Rated Readability
A	10	10	0	10.0
B	8	8	0	8.0
C	10	9	1	9.50
D	10	9	1	9.50
E	10	8.5	1.5	9.25

TABLE II

MEASURED READING ABILITIES, IN GRADE LEVELS, OF NINTH GRADE INDUSTRIAL ARTS STUDENTS

Reading Ability (Grade Level)	Number of Students	Percent of Total Group Tested
12.0 - 12.9	4	0.95
11.0 - 11.9	23	5.60
10.0 - 10.9	29	7.10
9.0 - 9.9	69	16.80
8.0 - 8.9	100	24.33
7.0 - 7.9	112	27.25
6.0 - 6.9	58	14.10
5.0 - 5.9	12	2.92
4.0 - 4.9	4	0.95
Total	411	100.00
M= 8.335 S.D. = 1.53		

up to his grade level would have scored approximately ninth grade level on the test of reading comprehension. However, as Table II indicates, 286, or nearly 70 percent, were reading below ninth grade level. Results of this testing program revealed a mean reading grade level for the group of 8.3. In fact, 186, or 45 percent, of the students had measured reading abilities of less than eighth grade level. To further describe this group it seems significant to point out that the Lorge-Thorndike Test of Intelligence was administered and that the data yielded a mean intelligence quotient of 100.9 with a standard deviation of 17.28. These quotients if plotted into a curve of distribution would closely resemble the normal curve of distribution.

Comparison of Readability with Reading Ability

An answer to the third basic question, which dealt with the relationship between textbook readaability and student reading abilities, was achieved by comparing the levels of textbook readability with the measured reading abilities of the ninth grade industrial arts students. This comparison reveals that a majority of the samples taken from Book A were rated above the measured reading ability level of 355, or 86.4 percent, of the ninth grade students tested. Book B presented a different picture since over 67 percent of the samples were within the reading abilities of a majority of the students in the group.

Book C presented a greater range of readability between samples within sections, as well as from section to section, than any other book. Less than one-half of the samples from Book C were rated within the reading ability levels of 50 percent of the ninth grade students. In fact, 20 percent of the samples taken from Book C were beyond the measured reading ability of 99 percent of the group.

The range of rated readability for samples within sections of Book D was smaller than the corresponding variability within sections of the other textbooks analyzed. Even though the amount of variability between samples was relatively small, 62 percent of the samples were rated beyond the measured reading levels of 86.4 percent of the ninth grade industrial arts students.

Book E contains 25 samples, 18, or 72 percent, of which were beyond the measured reading ability of 45.2 percent of the students tested. In fact 24 percent of the samples from Book E were rated higher than the reading abilities of at least 99 percent of the group.

Research Implications

Since this study revealed that some sections of general shop textbooks are too difficult for many ninth grade students to read, it is desirable for the industrial arts teacher to make greater use of information sheets, supplementary references, and other instructional materials that have a lower level of readability.

Publishers of textbooks for the general shop area of industrial arts should certainly be aware of the factors of readability, and they should assume the responsibility for providing reliable information regarding the readability levels of their textbooks. Such generalized terms as "controlled vocabulary" should be more completely defined in order to provide the teacher with a more specific index of readability.

In the preparation of textbook materials for the junior high level, authors should make an effort to control the factors of readability in order that the reading difficulty of the textbook might be more compatible with the reading abilities of a majority of students who will use them.

If the readability levels of these textbooks were compared with the reading abilities of students in the seventh and eighth grade, instead of the ninth grade, it can be assumed that the significance of these implications would be greatly amplified.

Implications for Teacher Application

The pattern of this research can be applied to any specific school situation. In fact, a measure of the reading abilities for students in industrial arts classes should be obtained by the industrial arts teacher in order that the instructional materials can be selected to more nearly suit the abilities of the students. The school counselor may have a measure of the reading grade placements for the students or he may be called upon to secure this basic information for the teacher. Since vocabulary is such an important factor in readability, it is desirable for the industrial arts teacher to alter his teaching procedures in order that more time may be given to the study of vocabulary. This is especially true in the areas of electricity and plastics.

It would seem desirable for the industrial arts teacher to also check the readability of his instructional material. It is the writer's feeling that a shortcut device, such as the "Reading Ease Calculator" produced by Science Research Associates, Inc., would be adequate for the classroom teacher's use in estimating reading difficulty. This calculator uses the same factors that are employed by Flesch in his Reading Ease formula.

Even though it is most advantageous to use a basic text or series of related references, it will become increasingly necessary for teachers to become more dependent on materials and procedures that bypass the textbook unless authors and publishers can meet this need for graded textbooks on the junior high school level.

FOOTNOTES

1. Wilbur R. Miller., Levels of Readability of General Shop Textbooks Compared with the Reading Abilities of Ninth Grade Industrial Arts Students, unpublished Ed. D dissertation, University of Missouri, 1960.
2. Rated readability and readability rating were used in the study when reference was made to an assumed grade level at which a piece of printed material can be read and understood, as indicated by an average of the grade levels obtained through the application of the Dale-Chall and Flesch formulae of readability.
3. Avis K. Porch, Reading Difficulty of Adopted Textbooks in the Social Studies for the State of Alabama, Grades III to XII, unpublished master's thesis, Alabama Polytechnic Institute, 1946).
4. Harry J. Siceluff, The Relative Difficulty of Noncontextual Vocabulary in Six Subject Matter Areas of the Ninth Grade, unpublished Ed. D dissertation, University of Missouri, 1948.
5. Jeanne S. Chall, Readability, An Appraisal of Research and Application, Monograph No.34
6. George W. Anderson, A Study of the Readability of General Business Training Textbooks, unpublished Ph. D. dissertation, University of Pittsburgh, 1955.
7. Nita Mae Wyatt, A Study of the Readability of Selected Social Studies Materials, unpublished master's thesis, University of Kansas, 1956.
8. E. F. Lindquist and A. N. Hieronymus, The Iowa Tests of Basic Skills - Form I, Boston: Houghton Mifflin Company.
9. Irving Lorge and R. L. Thorndike, The Lorge-Thorndike Test of Intelligence-Verbal Battery, Level 4, Boston: Houghton Mifflin Company.
10. Due to the structure of a "general shop" textbook which contains specific information from separate facets of industry such as drafting, woodwork, metalwork, electricity, etc., it was thought desirable to select at least three samples from each of these sections to get an indication of the relative reading difficulty of the content from the various sections. The results of this comparison should be regarded only as indications since the readability formulae were not designed to assess the readability of relatively small sections of a book but rather the book as a whole.

ERRATA

Re: June-July 1962 Journal of Educational Research, p. 21

We regret that the authors' names were not printed in the proper order: C. C. Walker should have preceded L. M. Stolurow, thus making the first footnote a more correct statement.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 4, December 1962)

Some Illustrated Comments on Hypothesis Construction And Research

DONALD J. WILLOWER
Pennsylvania State University

HYPOTHESIS CONSTRUCTION, theory development and research have an intimate but sometimes unclear relationship. This relationship has many facets, some of which may be illustrated by focusing on the problems dealt with by an investigator carrying out an empirical study based on theory. This paper reports and discusses certain problems encountered in a study guided by one of the current approaches to administrative theory.[1*]

The focus of this paper is upon "researcher behavior" rather than upon the content of the study. Several types of problems are considered. One deals with the selection of a problem for research and of a theoretical framework. Another type concerns hypothesis construction and the relationship between that activity and research. These problems are illustrated by reference to the study. While they grew out of a particular research activity, they appear to have quite general implications.

In order to provide a setting for further discussion, the theoretical framework which guided the study is briefly presented next.

The Theoretical Approach

The conceptual framework which was utilized was designed by Getzels and Guba,[2] and is a version of role theory. It postulates two general dimensions of social systems: the nomothetic dimension consisting of institution, role, and role expectation and the idiographic dimension consisting of individual, personality, and need-disposition. Social behavior is seen as a function of the interaction of the personality of a given individual defined in terms of need-dispositions with the requirements for a given role which the individual holds in a group or organizational setting.

Getzels and Guba have made several derivations from this theoretical framework which have special pertinence for administration. Those of major concern for our purposes involve leadership-followership styles. Three leadership styles were derived: the nomothetic, the idiographic, and the transactional.

In the nomothetic style, organizational requirements and conformity by subordinates to role expectations are emphasized. In the idiographic style, personality requirements and the tailoring of expectations to individual needs are stressed. In this style, expectations are kept relatively vague and informal. The transactional style is intermediate between the other styles and represents a combination of these styles. Because it has not yet been very precisely nor uniquely defined, it was omitted from the study under consideration.

It is of import for the hypothesis to be presented below to note that Getzels and Guba view the separate styles as different ways of "achieving the same goal; they are not different images of the goal."[3] The most expeditious route to goal achievement is seen, in the nomothetic style, as residing in the nature of the institutional structure and, in the idiographic style, as residing in the nature of the individuals involved.

The Broad Problem

The study in question grew out of a broad problem: the need for research based on theory.[4] Because of this need, one especially notable in studies in administration, it was decided that one of the administrative theories currently being developed would be used to derive a series of hypotheses and one of these hypotheses would then be tested empirically.

Thus, the study did not have its origins in a practical problem or in observed or supposed situational or behavioral incongruencies which seemed to call for examination and explanation.

---

[*] Footnotes will be found at the end of the article.

The approach used is not a common one and the thorough empiricist might argue that it is not a desirable one. However, it seems to the writer that one rather obvious way to make a contribution in a given field of study is to select a relatively adequate theory and test hypotheses derived from it. At any rate, the broad problem was selected as a starting point.

New problems for the investigator grew out of this decision. The most important of these was the question of which theory to use in deriving hypotheses. It seemed appropriate to restate this question to ask what criteria could be used to determine if a theory is an adequate one. A number of such criteria can be found in the literature. However, it must be recognized that their application is a matter of rational analysis and can give only a general indication of theoretical adequacy.

The criteria used were: (1) testability — a theory should lead to deductions which can be credited or discredited by empirical test; (2) consistency — a theory should be free from internal contradiction or be logically consistent; (3) scope — a theory should account for a relatively wide range of data; (4) predictive power — a theory should provide a basis for successful prediction; (5)fruitfulness — a theory should generate a variety of new hypotheses; (6) parsimony — other things being equal, a simpler theory is preferable to a more complicated one.

Of these, predictive power is perhaps the acid test of a theory. Here the Getzels-Guba framework possessed a decided advantage over some other formulations. Previous studies[5] by Getzels and Guba, Moyer, and Bidwell and Guba have been guided by that theory with some predictive success. This is especially notable in view of the general gap between theorizing and research found in present work in administration.

On the basis of the application of the criterion of predictive power and of the other criteria noted above, the Getzel-Guba theory was selected to guide the proposed research. It would be too time-consuming to detail the analysis involved in the application of each criterion to the Getzels and Guba framework here or to indicate why alternative approaches were not adopted. Two points should be made, however. The first is that the theory selected for use here is seen as one way of approaching administrative behavior. Other theories may have greater value and their development and utilization should, of course, be encouraged. On the other hand, and this is our second point, the criteria used appear to be useful in providing a general assessment of theory. This is especially true when theoretical formulations may be deficient in obvious respects, as has been undeniably the case with several approaches in administration.

Hypothesis Construction

In generating hypotheses from the Getzels-Guba theory, several problems had to be dealt with. One concerned the technique used to derive hypotheses. Another involved the question of the levels of abstraction of the hypotheses and their relation to evidence. A third grew out of the practical pressures which research exerts on hypothesis construction. Each of these problems will be considered separately.

Derivation of Hypotheses. The major problem here may be stated as follows: how should hypotheses be derived from a theory? If major emphasis is given to precision of logical connection, then perhaps hypotheses should be derived via symbolic logic, mathematically or at least syllogistically. Typically, however, these techniques are not used. In the most widely used approach, hypotheses are stated in ordinary prose and their rationale presented in the form of an argument. It may be that most social scientists feel more comfortable with that technique. On the other hand, many of our present theoretical concepts do not lend themselves readily to more rigorous treatment. For example, in the theory outlined earlier, we are not certain if the idiographic and nomothetic leadership styles represent opposite ends of an all-inclusive continuum or whether they and the transactional style do not account for all the possible variations of style along a hypothetical continuum. At any rate, the prose approach was used. The hypothesis which guided the study was:

> Administrators employing an idiographic leadership style will tend to regard subordinates as professionals to a greater extent than will administrators employing a nomothetic leadership style.

This hypothesis was derived as follows: in terms of the theory, idiographic leaders[6] see goal achievement taking place more readily when detailed role prescriptions are not imposed. Thus, subordinates are regarded as persons who, on their own, can and will behave so as to contribute to the achievement of organizational goals. The concept “professional” is introduced to designate this kind of subordinate, and is defined in terms of technical competence (can) and internalized high standards of occupational performance (will). Nomothetic leaders, regarding subordinates as persons whose effective operation must be insured by the imposition of detailed role prescriptions, would probably not regard subordinates as professionals to the same extent as would idiographic leaders. In deriving this hypothesis, the assumption was made that leaders generally employ a relatively consistent rationale for their leader behavior.

It should be noted that hypothesis construction as an activity may be conveniently differentiated from theory construction. This enables us, for example, to distinguish assumptions of the type above, which are involved in deriving hypotheses, from other more general assumptions which may be inherent in the nature of the underlying theory, in this case, role theory. When hypotheses are derived in the manner used here, it would seem necessary to clearly state the assumptions involved and to detail the rationale underlying their derivation. This is an obvious point but it is sometimes ignored.

Levels of Abstraction and Evidence. An aspect of the theory-research relationship needing clarity involved the level of abstraction at which a hypothesis is stated and its relation to evidence. Theoretical hypotheses are usually stated on relatively high levels of abstraction. They are usually tested via limited samples which are sub-classes of one of the variables linked in the hypotheses. For example, our hypothesis about "administrators" employing given leadership styles was tested via a sample of "school principals." In order to avoid attributing a greater degree of confirmation to the hypothesis than was warranted by the evidence gathered, a second hypothesis was stated on a lower level of abstraction as follows:

> Principals employing an idiographic leadership style will tend to regard teachers as professionals to a greater extent than will principals employing a nomothetic leadership style.

This hypothesis is readily derivable from our earlier hypothesis. The point to be noted is that evidence was gathered concerning the hypotheses stated on a lower level of abstraction and a less direct and less complete test is provided for the hypothesis stated on a higher level of abstraction.[7]

This essentially cautious device need not interfere with the rational and speculative operations necessary to hypothesis construction and theory development. It does have value in restraining extravagant confirmation claims for theoretical hypotheses and in pointing up further research needs. It might, for example, lead to research testing our hypothesis using a sample of leaders and subordinates in, say, an industrial setting.

Practical Pressures Exerted by Research. Hypotheses are stated as guides for research. Yet the research activity in turn exerts certain pressures on hypothesis construction. One of these is a pressure for what we will call "practical precision," or precision within the limitations of the resources of the investigator. Suppose we had stated the following hypotheses:

> Administrators employing an idiographic leadership style will tend to regard subordinates as professionals.
>
> Administrators employing a nomothetic leadership style will tend not to regard subordinates as professionals.

To test these hypotheses we would need to have a standard indicative of the perceptions of the professionalism of a given group of subordinates by a relatively large population. Because of pressures of time, budget and personnel, the development of such a standard would be a difficult task to accomplish properly. Thus, as a compromise, the more readily testable (and more easily derived) comparative hypothesis stated earlier was used to guide the research. This is not to say that there is something inherently wrong with comparative hypotheses. In fact, initial theoretical advances are often made by utilizing polar concepts, extremes and ideal types, that is, by focusing on comparisons. However, eventually scope must be increased and more precise measures found. General measures of long and short, for example, must eventually be replaced by readings in inches.

Concluding Comments

Scientific method is hardly a mechanical process. Creativity of conceptualization plays an integral part and researchers are often required to make unanticipated decisions and occasional compromises during the course of a study. Attention to the "research behavior" of the social scientist should provide further insight into the practical intricacies of scientific method. There would seem to be advantages for theorists and researchers inherent in the exchange of actual experience in carrying out their tasks.

This paper has reported some of the problems encountered in a study guided by theory. The way in which the problems were dealt with is indicated. Criteria for adequate theory were presented relative to the problem of selecting a theory to guide research. The problem of how to derive theoretical hypotheses was considered, as were problems relative to levels of abstraction and evidence, and the practical pressures exerted by research on hypotheses construction. Imperfections and inadequacies reflected in both the problems discussed and their tentative solutions are recognized. However, it is hoped that their recitation will have some value for others engaged in similar activities.

FOOTNOTES

1. Administration is a field which has shown a concern for systematic theory only in recent years. In educational administration, see Andrew W. Halpin (ed.), Administrative Theory in Education (Chicago: Midwest Administration Center, University of Chicago, 1958), and Daniel E. Griffiths, Administrative Theory (New York: Appleton-Century-Crofts, Inc. 1959). An earlier publication of some importance is Arthur P. Colardarci and Jacob W. Getzels, The Use of Theory in Educational Administration (Stanford, California: Stanford University Press, 1955). In business administration see, for example, Chris Argyris, Personality and Organization (New York: Harper and Brothers, 1950). In public administration Herbert A. Simon and Dwight Waldo have both been concerned with administrative theory. The Administrative Science Quarterly, published by the Graduate School of Business and Public Administration at Cornell University, has given considerable attention to theoretical problems in administration since its inception in 1956.
2. J. W. Getzels and E. G. Guba, "Social Behavior and the Administrative Process," The School Review, LXV (Winter, 1957), pp. 423-41.
3. Ibid., p. 435. Italics in original.
4. See Daniel E. Griffiths, "Toward a Theory of Administrative Behavior," Administrative Behavior in Education, Roald F. Campbell and Russell T. Gregg (eds.), (New York: Harper and Brothers, 1957); Coladarci and Getzels, op. cit.; and John Walton, "The Theoretical Study of Educational Administration," Harvard Educational Review, XXV (Summer, 1955), pp. 169-178.
5. See Jacob W. Getzels and Egon G. Guba, "Role, Role Conflict, and Effectiveness: An Empirical Study," American Sociological Review, XIX (April, 1954), pp. 164-175; Donald C. Moyer, "Teachers' Attitudes Toward Leadership as They Relate to Teacher Satisfaction," unpublished doctoral dissertation (University of Chicago, 1954); and Egon G. Guba and Charles E. Bidwell, Administrative Relationships (Chicago: Midwest Administration Center, University of Chicago, 1957).
6. For purposes of the study, it was assumed that those in superordinate administrative positions in a formal organization are "leaders" with respect to a subordinate group or person.
7. The design, methodology and results of the study used as an example are not central to our present purpose and are reported elsewhere. See Donald J. Willower, "Leadership Styles and Leaders' Perceptions of Subordinates," Journal of Educational Sociology, XXXIV (October, 1960), pp. 58-64.

ERRATA

We are very sorry for a serious error in the Klausmeier-Check article published in the April 1962 Journal of Educational Research, page 322. The last sentence in the next to last paragraph was left unfinished and the sub-title "Summary" was omitted. The correction is:

> Upon ascertaining this and knowing learnings at next higher levels of difficulty, the teacher should offer individualized instruction or arrange the children in smaller groups as necessary to provide instruction which will encourage efficient acquisition and subsequent retention and transfer of learning.
>
> Summary
>
> Children of low (55-80 IQ), average (90-110 IQ), and high intelligence. . . etc.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 4, December 1962)

Differences in Achievement Test Scores Associated with Item Exposure Time and Audio Reinforcement*

R. P. KROPP, F. J. KING
W. L. BURR, J. G. BEARD
Florida State University

This report deals with the administration of an achievement test by projecting the test items on a screen instead of presenting them in booklet form, the usual paper-and-pencil mode of administration. The possibility of testing in this manner is worthy of study at this time for a number of reasons including the advent of mass instruction via television. Administering a test by projecting the items creates restrictions and problems; e. g., only one or a small group of items can be presented at the same time, the examinee cannot refer back to previously exposed items, exposure times must be predetermined, and the method is new and strange to test-booklet-oriented students. The study reported here deals with two problems, exposure time and the audio reinforcement of the projected items.

Problem

The specific problems dealt with are:

1. Whether there is a differential effect on achievement test scores between (a) exposing items for a constant time and (b) exposing items for variable times on the basis of reading load.
2. Whether there is a differential effect on achievement test scores between (a) projected, i. e., visual presentation of items on a screen and (b) an audio-visual presentation of the items, i. e., having the stem and alternatives read aloud to all examinees when the item appears on the screen.
3. Whether there is an interaction effect between (a) constant-variable exposure time and (b) projected visual- and projected audio-visual presentation of items.

In addition, data were analyzed according to mental ability levels of examinees to determine if experimental variables affect various ability levels differentially.

Procedures

Test: The test used for experimental purposes was a fifty-item multiple-choice test on the content of American Government which is taught at the lower undergraduate level. The test was constructed by the instructor in the course by randomly selecting items from a pool of items that were discriminative on previous administrations.

Items were sorted into three categories according to the number of words in the stem and alternatives of each item. Thus, reading load was construed to be "reading amount." Half the items in each category were assigned numbers from 1-25; the other half, from 26-50. Thereby, each half of the test was comparable on reading load. However, since item difficulty data were not available to the investigators, it was not known in advance whether the half-tests were of comparable average difficulty.

For the normal administration, the fifty-item test was prepared in the usual way--course title, directions, sample items, and the test items. For the experimental administrations, each item was reproduced on a 2 x 2 slide, one item to a slide.

Subjects: The subjects were sixty-four students registered in two sections of American Government which were being taught by the same instructor. Each group had been taught approximately the same content. American Council on Education Psychological examination scores were available for most of the students. Students in each section, separately, were sorted into control and experimental groups, roughly in the proportion of one to two, on the basis of ACE scores. Students for whom scores were unavailable were sorted into the two groups.

Testing conditions: The control group of each section was administered the test under usual

---

* This study was supported by a grant from the United States Office of Education, National Defense Education Act of 1958, Title VII, Grant Number 7-08-075.

conditions by the course instructor. The test time was forty-eight minutes, which proved to be ample.

One experimental group created by dividing one of the sections was administered the test under the following conditions: each of items 1-25 was given a constant exposure, 42 seconds; items 26-50 were given variable exposures--32, 42, and 52 seconds-according to reading load; and for all items the projected item was read to the students. The other experimental group was administered the test under the following conditions: items 1-25 were given variable exposure times as indicated above; each of items 26-50 was given a constant exposure time of 42 seconds; and there was no audio reinforcement of the projected items.

The exposure times mentioned above were calculated in the following ways. Previous studies by the investigators indicated that when a test is administered by projecting it on a screen the test taking time can be reduced substantially from what is needed when the same test is administered in booklet form. On the basis of their experience they decided that the fifty-item achievement test used here would require thirty-five minutes, at most, when administered by the projected method. The constant exposure time of 42 seconds was equivalent to giving each of the fifty items its proportionate amount of the thirty-five minutes. The variable exposure times were determined by giving 42 seconds to the middle category of reading load, 10 seconds less to items having the least reading load, and 10 seconds more to items having the highest reading load.

The audio reinforcement of the projected items was accomplished by electronically synchronizing a tape recorder and slide projector such that the tape recorder at predetermined intervals triggered the automatic changer of the projector. The course instructor taped the audio portion. Thus, the moment an item appeared on the screen the instructor's reading of the entire item was broadcasted.

Items were projected onto a 3' x 5' screen. All projected items could be seen clearly by subjects; the audio presentation was sufficiently amplified to permit comfortable auditing. There was no inter-exposure interval of any consequence inasmuch as slide changing was almost instantaneous. Subjects were given all directions, sample items, etc., by projection. The directions contained information about the length of the constant exposure interval and the variable exposure intervals and which block of items would receive constant exposures and which would recieve variable exposures; however, students were not informed of the exposure time for a particular item when it appeared on the screen. The investigators administered the test under normal conditions.

Analysis of data: The following analyses of the data were executed.

1. Analysis of variance on control group scores on items 1-25, and 26-50 separately, to determine differences between half-tests and control groups.
2. Analysis of variance of scores from control group (control group from each section combined) and from both experimental groups on each half-test, separately, and for the whole test.
3. Analysis of variance of scores from experimental groups according to testing condition, item exposure time, and level of mental ability of subjects.

Results

Analysis of variance between control groups and half-tests. The results of the analysis of variance between control groups and half-tests appear in Table I. They indicate that significant differences occurred between half-tests. The interaction between groups and half-tests was significant. The means appear in Table II. The data in Table II indicate control group A was superior, but not significantly so, to group B, and that items 26-50 were less difficult than items 1-25. Perhaps the significant interaction is due to differences in the instruction of the two groups.

Analysis of variance of control and experimental groups on each half-tests and the whole test. For comparison purposes the two control groups were lumped into a single control group. Hereafter, control group refers to this combination. The analysis of variance data for the three groups on each half-test, separately, and the whole test appear in Table III. No significant between group differences were noted in either half-test or the whole test.

Analysis of variance of scores from experimental groups by testing condition, item exposure time, and level of mental ability of subjects. Although previously cited analyses failed to reveal between group differences, inspection of the score distributions revealed substantial differences between groups. Consequently, the analysis of variance in experimental conditions was done by ACE level to rid group variation of that variation due to subject differences in mental ability. Each experimental group consisted of five high-scorers on ACE and ten low-scorers. Each group consisted of matched subjects. The results of the analysis appear in Table V. The data reveal that no differences were occasioned by the projected audio-visual modes of presenting test items.Item exposure time of a constant interval and a variable interval did not lead to significant differences in scores. When data were analyzed according to level of intelligence a significant difference on the criterion variable was noted and it was expected. The only significant interaction occurs between mode of presentation and item exposure time. Because the half-tests were not of comparable difficulty, the interaction should be regarded as spurious. Unequivocal interpretations of the data are difficult to make owing to the lack of comparability of half-tests. Inspection of the data leads the investigators to believe that on items of constant exposure time the projected audio condition produces higher scores than does projected alone. Also, it seems for high scoring students on the ACE the projected audio condition is superior to the projected condition alone on items of variable exposure time; however, this seeming difference does not hold for students having low ACE scores.

Summary

Groups of college freshmen and sophomores who were registered in an American Government course were administered a fifty-item multiple-choice test constructed by the course instructor. One group was administered the test under normal conditions. Another group took the test visually in the sense that each item, one item at a time, was projected onto a screen which served as the only exposure of that item to the student. The third group took the test under projected conditions but in addition when the item appeared on the screen it was read to them by means of a tape recording. Under the projected conditions, a block of twenty-five items received a constant exposure interval and the remaining items received variable exposures in accordance with the reading load of each item. The order of blocks was different for the two groups; thus, one group had constant exposure intervals on items 1-25, whereas the other group had constant exposure intervals on items 26-50. The analysis of data revealed that no between group (one control and two experimental) differences occurred on either half-test or the whole test. However, analysis of the control group data revealed that the half-tests were not of comparable difficulty. Further analyses were confined to the two experimental groups. The analyses revealed no significant between group differences could be attributed to mode of presentation (projected, and projected audio) or to item exposure time (constant and variable). However, the interaction of these variables was significant but it was regarded as spurious due to the lack of comparably difficult half-tests. There was some evidence that the projected audio condition is superior to the projected condition alone on items of constant exposure time.

Generally the data support the notion that the administration of tests by projecting items onto a screen to be viewed and responded to by students produces scores similar to those obtained under normal administrative conditions, thereby lending support to the feasibility of administering tests by television.

TABLE I

ANALYSIS OF CONTROL GROUPS BY HALF-TESTS

Source of Variation	df	ss	ms	F	p
Between Groups	1	96.95	96.95	3.64	>.05
Error$_B$	23	612.53	26.63		
Between Halves	1	40.50		12.54	<.01
G x H	1	18.75		5.80	<.05
Error$_W$	23	74.25	3.23		

TABLE II

MEAN SCORES OF CONTROL GROUPS ON HALF-TESTS

Control Groups	Half-Tests 1	Half-Tests 2
A (N=14)	12.86	13.57
B (N=11)	8.82	12.00

TABLE III

ANALYSIS OF VARIANCE OF GROUPS ON HALF-TESTS AND WHOLE TESTS

Test	Source of Variation	df	ss	ms	F
First-Half	Groups	2	13.91	6.96	
	Error	61	808.53	13.25	
Second-Half	Groups	2	18.29	9.15	
	Error	61	901.57	14.78	
Whole Test	Groups	2	63.51	31.76	
	Error	61	2,782.85	45.62	

TABLE IV

GROUP MEANS ON EACH HALF-TEST AND WHOLE TEST

	Groups		
Test	C (N=25)	X_A (N=15)	X_B (N=24)
First-Half	11.08	12.13	11.96
Second-Half	12.88	13.93	14.00
Total Test	23.96	26.06	25.96

TABLE V

ANALYSIS OF VARIANCE OF SCORES FROM EXPERIMENTAL GROUPS

Source of Variation	df	ss	ms	F	p
Presentation (P)	1	26.66	26.66	1.87	>.05
ACE (A)	1	66.01	66.01	4.64	<.05
P x A	1	4.41	4.41		
Error B	26	369.85	14.22		
Item Exposure (I)	1	.26	.26		
P x I	1	41.67	41.67	5.81	<.05
A x I	1	5.21	5.21		
P x A x I	1	1.41	1.41		
Error W	26	186.45	7.17		

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 4, December 1962)

A Formula for Predicting the Comprehension Level of Material to be Presented Orally

JOHN R. ROGERS
University of Mississippi

THE ABILITY to communicate is undoubtedly one of man's most valuable attributes. Communication may be simply defined as a process which allows an idea to be transmitted from one mind to another. Essential to the process are: (1) a mind to formulate the idea, (2) a method for transmitting the idea, (3) a mind to receive the transmission, to interpret the stimuli which it involves, and to recreate the idea in a form as nearly identical to the original as possible.

Man has developed many systems for the carrying on of this process. There are pictures, songs, statues, written words, spoken words, eyebrows to be lifted, and shoulders to be shrugged. Of the many avenues available to him for communicating ideas, there is little doubt that the spoken word is most used by man.

There are broadly discernable stages in linguistic development. Language development may be thought of as a continuum, one extreme of which is characterized by a limited vocabulary and a simple grammatical structure, and the other extreme of which is characterized by an extensive vocabulary and a complex grammatical structure. If this assumption is correct it should be possible to devise a method for predicting the general stage of language development which would be required by an individual if he is to understand any specific body of material which is to be presented orally.

A method for predicting the comprehension level of material to be presented orally would be of value to many people. Lawyers, radio broadcasters, public speakers, preachers, and teachers would all welcome such a technique. As the average educational level of adult Americans is somewhere around the ninth grade, it would be futile for a radio announcer to advertise his sponsor's product in terms which could be understood only by a superior college graduate. Similarly, a teacher who "explains" to a third grade class in phrases which are unintelligible even to average children in the eighth grade is wasting not only his own time but that of his charges as well. What is needed is a technique whereby a speaker can grade or evaluate material in terms of the comprehension level of those who are to experience its oral presentation.

It is immediately apparent that the general problem of oral comprehension may be divided into two major sub-areas: (1) content factors, (2) delivery factors. The first of these deals with elements which are found in the material itself--variety and difficulty of vocabulary, grammatical complexity and familiarity and interest of the general subject. The second deals with elements which depend primarily upon the person who delivers the material—dialect, enunciation, rate of delivery, and quality of the speaker's voice.

As the scope of this article must be restricted to manageable proportions, its attention is directed toward factors which are inherent in material *per se*. It undertakes to describe the derivation of a formula for predicting the comprehension level of materials to be presented orally. This formula is based on content factors and their relationships to the comprehension levels of specific audiences.

The most logical method for arriving at the desired formula was obvious from the beginning: (1) identify a number of content factors which might influence the comprehension level of material to be presented orally, (2) as a criterion identify a body of material graded according to its understandability by people at all stages of language development, (3) through the use of partial and multiple correlation technique, using the criterion proposed in (2) above as the dependent variable and the proposed measures of grammatical complexity and vocabulary loading as independent variables produce a formula which would predict the level of language development required for comprehending any specific body of material to be presented orally.

As a matter of fact, that was the actual route followed in deriving the final formula. Many sticky problems were encountered on the way, however. There is not space to list them here, and a description of them and their solutions would add little light to this presentation.

The identification of the content factors to serve as independent variables was a simple matter. As grammatical complexity is a function structure rather than a function mode of delivery, content factors which had been found significantly related to difficulty in printed material might well be assumed to play a similar role with respect to material for oral presentation. As there was in existance a rather extensive literature pertaining to readability, a thorough study of this material was indicated. An analysis of each of twenty-one formulae for predicting readability revealed eight variables which were significantly related to readability: (1) average sentence length, (2) subordination index, (3) average clause length, (4) average number of prepositional phrases per 100 words, (5) average number of infinitive phrases per 100 words, (6) average number of both infinitive and prepositional phrases per 100 words, (7) average number of words per 100 not on Dale's longer list,[1] (8) average number of words not on Dale's shorter list.[2] With the exception of (2), "subordination index," the names of the eight variables are self-explanatory. Subordination index is determined by adding the number of dependent clauses in a passage to the number of independent clauses and dividing the number of dependent clauses by the resulting total.

Average sentence length presented some difficulties. Although a high correlation between average sentence length and readability had been revealed by a majority of the studies relating to the problem:

> ... it is impossible to determine what constitutes a sentence in an individuals' oral or written composition, unless the sentence be perfectly punctuated by marks or inflection. In spoken composition inflection frequently is not a guide.[3]

For reasons which are made apparent in the descriptions of the dependent variable to follow, it was not feasible to use average sentence length as one of the independent variables. Instead, a concept which for want of a better name may be designated the idea unit was developed. This neologism was coined to designate a proposed new measure of linguistic complexity which may be thought of as falling on an imaginary grammatical scale somewhere between the clause and the sentence. The "idea unit" is used to designate an independent clause, together with its modifying elements.

It is suggested as a more accurate measure of linguistic complexity than average sentence length for several reasons. Although a compound sentence containing three independent clauses would logically appear to be little more difficult to understand than would three simple sentences, when sentence length is used as a measure in such cases, the compound sentence is judged to be considerably more difficult. Furthermore, a compound sentence containing four independent clauses and thirty words is judged to be equally as difficult as a complex sentence containing one independent clause, three dependent clauses, and an equal number of words. This seems implausible. Obviously, a number of simple sentences strung together by conjunctions to make a long compound sentence would be less difficult than a complex sentence of equal length. Unlike sentence length, which ignores internal structure and considers number of words only, if idea unit length is used as a measure, the complex sentence is judged considerably more difficult than the compound sentence.

The independent variables selected, together with their respective modes of appraisal, were as follows:

> The idea unit length is determined by dividing the number of words in the transcript by the number of independent clauses.
>
> Subordination index is determined by adding the number of dependent clauses to the number of independent clauses and dividing the number of dependent clauses by the resulting total.
>
> Average clause lenth is determined by dividing the number of words by the number of clauses.
>
> The average number of prepositional phrases per 100 words is determined by dividing 100 by the number of prepositional phrases in a hundred word sampling.
>
> The average number of infinitive phrases per 100 words is determined by dividing 100 by the number of infinitive phrases in a hundred word sampling.
>
> The average number of both infinitive and prepositional phrases per 100 words is determined by dividing 100 by the number of both infinitive phrases and prepositional phrases in a hundred word sampling.
>
> The average number of words per 100 not occurring in the Dale short list is determined by dividing 100 by the number of such words in a hundred word sampling.
>
> The average number of words per 100 not occurring in the Dale long list is determined by dividing 100 by the number of such words in a hundred word sampling.

In order that the proposed hypothesis might be tested, one further thing was lacking--an adequate body of material accurately graded according to its difficulty of comprehension when presented orally. An oral counterpart of the McCall, Crabbs Standard Test Lessons in Reading would have been ideal. Unfortunately, no such body of material existed, and the production of a McCall, Crabbs type of criterion would have been a major project in itself; but a way around the dilemma was found . . .

Although we can never be quite sure that one person comprehens the oral language of another, we can safely assume that a speaker normally understands his own language. It also seems probable that although a typical five year old may not understand the language of an adult, he is quite likely to understand the language of another five year old. In other words, it was assumed that a child understands his own lanuage and that of other children like him. As a result of this assumption the experimental design virtually outlined itself.

(1) collect language samples from a large number of typical children from all age-grade levels;

(2) analyze these language samples in terms of the eight content factors previously described;

(3) though the use of partial and multiple correlation technique, using grade level as the dependent variable, work out a prediction formula for applying to any material to be graded for listenability.

These things were done. Individual and separate samples of spontaneous and unrehearsed language were recorded from at least eight children in each grade of five elementary schools, five junior high schools, and five senior high schools--480 samples in all. Each school in the respective divisions was in a different city. In order to insure that the criterion be composed of the speech of "typical" children, each child selected was one who:

(1) was approximately as old as a child would be if he had entered the first grade at seventy-eight months of age and had neither failed nor skipped a grade;

(2) had a derived I. Q. between 90 and 110;

(3) came from a home in which only English was spoken;

(4) represented neither extreme of the socio-economic scale;

(5) was considered an average child by his teachers.

Four hundred and eighty recordings of more than a hundred words each were used in developing the criterion. An exact typed reproduction of each recording was made. In effect, these transcriptions represented twelve bodies of material--one for each of the twelve school grades

These samples were analyzed in terms of the eight independent variables previously identified. Using the Wherry-Doolittle test selection method (which identifies the variables which yield a maximum R with the criterion) the following regression equation was worked out:

$$G = .669\,I + .4981\,LD - 2.0625$$

in which G = grade level, I = average idea unit length, and L.D. = the average number of words in a hundred word sampling which do not appear on Dale's long list. The - 2.0625 is a constant.

The multiple correlation coefficient of G with I and LD was found to be .727, and the standard error of the equation's estimate was 2.3749; the G predicted by the equation would be accurate within a range of about two grades above or below the predicted one more than two-thirds of the time.

From the nature of the procedures followed in deriving the formula, it is obvious that the last word has not yet been spoken. Furthermore, the formula is presented as untried, untested, and unproved. At present there is no graded body of material against which it may be tested. Perhaps other researchers will work out methods for testing the formula in practical situations.

FOOTNOTES

1. Edgar Dale and Jeanee S. Chall, "A Formula for Predicting Readability: Instructions," Educational Research Bulletin, 27: 37-54, p. 50, February 17, 1948.
2. Edgar Dale, " A Comparison of Two Word Lists, " Educational Research Bulletin, 10: 484-89, p. 118, December 1931.
3. Lou L. LaBrant, "A Study of Certain Language Developments of Children in Grades Four to Twelve Inclusive, " Genetic Psychology Monographs, 14: 387-491, p. 114, 1933.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 4, December 1962)

The Effect on Typewriting Performance of Varying the Degree Of Interest in Timed-Writing Copy Materials

WILLARD H. NELSON; BETSY ANN LITTLE; F. J. KING
Florida State University

THE PURPOSE of this study was to determine whether interest value of copy materials for timed writing affects typing speed and/or accuracy. Discussions of desirable practices (Dvorak, et al., 1936; Fuller, 1943; Tidwell and Bell, 1951; Budish, 1954), have indicated that teachers of typewriting think that interesting copy promotes attention, speed, and accuracy. Research findings interpreted to support this contention were reported by Bell (1950) and McQuade (1957). However, statistically non-significant findings favoring uninteresting copy were reported by Dodson (1959), while non-significant figures in the opposite direction were found by Alexander (1952). Both of these investigators used high-school students as subjects.

In view of the fact that strong opinion on the subject has remained unsupported by statistical evidence, it appeared desirable to test the null hypothesis as applied to the speed scores and to the accuracy scores obtained on copy which had been rated high and copy which had been rated low in interest value.

Methodology

Subjects

Subjects in the present study were 37 female and 14 male students enrolled in first year typing classes at the Florida State University. The students were in three sections taught by the same teacher.

Materials

The procedure for selecting "interesting" and "non-interesting" copy differed from that of previous investigators who had used textbook materials (Dodson, 1959; Alexander, 1952). Copy for this study was taken from non-textbook sources because of the possibility that they would provide a different range of interest value and readability from that in textbooks. In order to control both interest level and typewritability of the material and to obtain substantial differences in interest value, the following procedures were used to select materials.

(1) Typewritability of diverse materials was assessed according to Bell's C scale for typewritability per 100 running words (Bell, 1950). The formula, which takes into account stroke intensity (X_1), syllabic intensity (X_2), and percentage of frequently used words (X_3), is $C = -2.4370X_1 + 3.4945X_2 + .0844X_3 + 7.8144$. Twelve selections which were homogenous with respect to typewritability at a median level of difficulty (C value between -.9055 and .0862), were selected for further examination.

(2) Interest value of each of the twelve selections was assessed through a rank ordering by college classes of typing students. The students placed the copy sheets in a pile in rank order. These same students also rated each copy separately on a 7-point scale ranging from very unin-

teresting (7) to very interesting (1). Markedly different reactions of different individuals to the same copy led to the elimination of several possible selections. Marked differences between the sexes accounted for other eliminations; the final selections were different for men and women. The mean scale value of the women's "interesting" copy was 2.45, of the men's, 1.94; the uninteresting were 5.85 and 5.65, respectively.

The resulting C values on Bell's scale of typewritability for the interesting and uninteresting copy were -.26 and -.28 for the women's selections and -.79 and -.78 for the men's selections.

Procedure

The order of the presentation of the copy material was such that one half of each sex group typed interesting materials first and one half typed the non-interesting first. Within each sex group, the order of typing was randomly determined. The classroom teacher, who was carefully initiated into the project and who had previously obtained the interest ratings, administered the copy in standard 2-minute timed typewriting tests, following her regular procedures for such exercises.

Speed and accuracy scores were obtained for each subject on each copy. Gross words per minute and total errors were calculated in accordance with International Typewriting Contest Rules.

Results and Discussion

The Mean of gross words per minute on interesting material was 31.76; for uninteresting material it was also 31.76. The mean number of errors on interesting material was 4.29; on uninteresting material it was 4.61. The obtained t value for the difference between means for errors was .74.

The sustaining of the null hypotheses, i.e., of non-significant differences in speed and accuracy of typing on interesting and uninteresting copy material, supports the non-significant differences in the Dodson and Alexander studies. It also tends to contradict Bell's contention that dull material leads to slower typing; however, the present study, being essentially cross-sectional in nature, does not preclude the development of cumulative differential effects.

REFERENCES

1. Alexander, Margaret Ann. "The Effect of the Degree of Interest in Straight Copy Material on Pupil Performance in Typewriting," Unpublished Master's Thesis (University of Washington, 1952).
2. Bell, Mary LaVerne. "Some Factors of Difficulty in Typewriting Copy," Unpublished Ed.D. Dissertation (University of Oklahoma, 1950).
3. Budish, Bernard E. "An Evaluation of Textbooks in Business Education," Journal of Business Education, XXX (1954), pp. 68-70.
4. Dodson, Glenna A. "The Effect of Interesting and Non-interesting Copy Material on Speed and Accuracy in Typewriting," Unpublished Ed.D. Dissertation (University of Florida, 1959).
5. Dvorak, M., et al. Typewriting Behavior (New York: American Book Company, 1936).
6. Fuller, Donald C. "Causes of Errors in Typewriting," Journal of Business Education, XIX (1943), pp. 19-21.
7. McQuade, Sister M. Edward. "The Role of Motivation Techniques in Vitalizing Typewriting," Unpublished Master's Thesis (Kansas State Teachers College, 1957).
8. Tidwell, M. Fred, and Bell, Mary L. Tested Timed Writings (New York: Prentice-Hall, 1951).

Field News

THE CENTER FOR PROGRAMED INSTRUCTION CONDUCTED two surveys for the U. S. Office of Education to determine the availability and usage of Programed Instruction. Educators and psychologists should find the two resulting publications of interest and value in teaching and research.

The first questionnaire went to all known publishers and suppliers of progamed material. Sample frames of each program available for school use in September 1962, were requested.

A second questionnaire was sent to 15,000 superintendents across the country to determine their knowledge about, interest in, and actual usage of all forms of programed instructional materials.

Both publications, the listing of available programs—designated as Programs, '62—and the survey of utilization—The Use of Programed Instruction—may be obtained from the U. S. Government Printing Office.

Both may be obtained from the Superintendent of Documents, Washington, D. C. Programs, '62, OE 34015 is $1.50. “Usage” should be available in November 1962, and will cost approximately $1.00.

THE UNIVERSITY COUNCIL FOR EDUCATIONAL ADMINISTRATION proposes a five-year plan for the period 1964 to 1969 which will be oriented toward research development. The planned activities include a series of research seminars in a number of substantive areas of administration, a task force to study the development of graduate programs to train researchers in educational administration, a task force to consider and report on the post-doctoral development of researchers, cooperative research task forces and task forces located at individual universities, the development of proposals for funding to the National Institutes of Mental Health and the Cooperative Research Programs, and the development of abstracts of completed and on-going research.

Regional meetings will be held. Representatives from all of the UCEA institutions can plan their participation in these activities.

ONE HUNDRED AND SIXTY-NINE RECORDED television courses are available for broadcast in the United States. Of these 65 are for use in elementary school, 57 for college-level instruction, five for in-service teacher education, and seven for adult education. These courses may be obtained from 41 different educational producers.

Information about these courses is contained in “Instructional Television Materials: A Guide to Films, Kinescopes, and Videotapes Available for Televised Use”, just published by the National Instructional Library in conjunction with its two regional counterparts in Lincoln, Nebraska, and Cambridge, Massachusetts.

RELIGIOUS AFFILIATION IS RELATED TO PARENTS' IDEAS about themselves and their children. Professor Paul H. Whiteman, director of the University of Wisconsin Psycho-Educational Clinic, conducted a study of 30 parent couples from three denominations: Baptist, Jewish, and Roman Catholic. He found that personality and attitude tests revealed surprising similarities between Catholic and Baptist parents and some striking differences between these groups and Jewish parents. For example, he found that Jewish parents encourage their children to speak up more than do Catholic and Baptist parents. This, he feels, may be related to the less rigid authority structure characteristic of present-day Judaism as compared with the other two religions. He found that Jewish parents believe that a child's ideas should be seriously considered in making family decisions, but that Roman Catholic or conservative Baptist parents are less likely to take the child seriously in making such decisions.

TO ENCOURAGE THE EXPANSION OF SYSTEMIZED SCIENTIFIC KNOWLEDGE in the field of television communication, eight social scientists and educators, together with Television Bureau of Advertising, launched a competition for exceptional plans in the field of television research. The competition was limited to plans rather than to accomplished research because, in the opinion of the governing committee, the definition of the research problem and the design of the research strategy are the more critical phases of any research.

The competition was set very broadly: research plans could be directed to problems referring to any aspect of human behavior and television, television by itself, or television in relation to other means of communication. The closing date for submitting plans was set for March 15, 1961. One hundred and forty-seven entries were received. More than half were from faculties and staffs of 59 different colleges and universities. Of all the plans submitted, 18 were selected for reward.

THE COMMITTEE FOR ECONOMIC DEVELOPMENT, a group of 200 businessmen and educators, has helped to shape and influence many public and private policies for almost a generation. The Committee has employed objective research and discussion to devise policies that will promote economic growth, high employment, and stable prices. Up to now the results of the Committee's work has been available to a limited number of business leaders, scholars, and decision makers. The growing interest in national economic health and its close interrelationships abroad make it important that the carefully researched data of the CED trustees become available to an ever increasing audience. Persons with interest in social and economic development now may share in the results of the work of the trustees by writing to John H. Crider, Deputy Director of Information, Committee for Economic Development, 711 Fifth Avenue, New York 22, New York.

•••

Readers are invited to send items of research interest to John Schmid, Professor, College of Education, University of Arkansas, Fayetteville, Arkansas.

Editorial Board

JACOB O. Bach, Head, Educational Research Bureau, Southern Illinois University, Carbondale, Illinois. Critique Editor.
WILLIAM A. BROWNELL, Dean, School of Education, University of California, Berkeley, California. Education Psychology, Learning and Teaching.
ROBERT A. DAVIS, Professor of Educational Research, George Peabody College for Teachers, Nashville, Tennessee. Evaluative Techniques.
GUY D. McGRATH, Dean, College of Education, Arizona State University, Tempe, Arizona. Secondary Education.
HAROLD E. MITZEL, Assistant Dean for Research, Pennsylvania State University, University Park, Pennsylvania.
WILLARD C. OLSON, Dean of College Education, University of Michigan, Ann Arbor, Michigan. Child Development.
JOHN SCHMID, Professor, College of Education, University of Arkansas, Fayetteville, Arkansas. Field News Editor.
WILSON B. THIEDE, Professor of Education, University of Wisconsin, Madison, Wisconsin. Book Review Editor.
ARTHUR E. TRAXLER, Executive Director, Educational Records Bureau, New York 32, New York. Educational Measurements.

Associate Editors

HAROLD M. ANDERSON, Associate Professor of Education, University of Colorado, Boulder, Colorado.
VERNON E. ANDERSON, Dean, College of Education, University of Maryland, College Park, Maryland.
RUDYARD K. BENT, Professor of Education, University of Arkansas, Fayetteville, Arkansas.
EMMETT ALBERT BETTS, Research Professor, University of Miami, Coral Gables, Florida
LAWRENCE P. BLUM, Assistant Director, Student Personnel Services, University of Wisconsin-Milwaukee, Milwaukee 3, Wisconsin.
J. MURRAY LEE, Chairman, Department of Elementary Education, Southern Illinois University, Carbondale, Ill.
D. WELTY LEFEVER, Professor of Education, University of Southern California, Los Angeles, California.
L. JOSEPH LINS, Professor and Coordinator, Institutional Studies, University of Wisconsin, Madison, Wisconsin.
S. E. TORSTEN LUND, Associate Professor of Education, University of California, Berkeley, California.
WILLIAM A. McCALL, Lecturer in Education, University of Miami, Coral Gables, Florida.
DELYTE W. MORRIS, President, Southern Illinois University, Carbondale, Illinois.
VALWORTH PLUMB, Chairman, Division of Education and Psychology, University of Minnesota (Duluth Branch), Duluth, Minnesota.
DAVID G. RYANS, Rand Corporation, Santa Monica, California.
LESTER B. SANDS, Professor of Education, University of California, Santa Barbara, Goleta, California.
MANDEL SHERMAN, President, and Consultant, Emotional Health Association, Los Angeles, California.
MILDRED B. STANTON, Consultant, Special Education, State Department of Education, Hartford 15, Conn.
W. W. THEISEN, Professor of Education, Marquette University, Milwaukee, Wisconsin.
M. R. TRABUE, Professor of Higher Education, University of Kentucky, Lexington, Kentucky.
J. WAYNE WRIGHTSTONE, Director, Bureau of Educational Research, Board of Education of City of New York, Brooklyn, New York.

Editors Emeriti

E. J. ASHBAUGH, Dean Emeritus, School of Education, Miami University, Oxford MR50, Ohio.
HARRY J. BAKER, 19050 Wiltshire, Lathrup Village, Michigan.
FOWLER D. BROOKS, Director Emeritus, Department of Psychology, DePauw University, Greencastle, Ind.
LEO J. BRUECKNER, Professor Emeritus, University of Minnesota, Minneapolis, Minnesota.
B. R. BUCKINGHAM, 33 Chatham Circle, Wellesley Hills, Massachusetts.
EDGAR A. DOLL, Director Emeritus, Devereux Schools, Devon, Pennsylvania.
HARL R. DOUGLASS, Director Emeritus, School of Education, University of Colorado, Boulder, Colorado.
E. DUNCAN GRIZZELL, Dean Emeritus, University of Pennsylvania, 640 Maxwelton Court, Lexington, Ky.
ERNEST HORN, Professor Emeritus of Education, State University of Iowa, W307 East Hall, Iowa City, Iowa.
EARL HUDELSON, Professor Emeritus of Education, West Virginia University, Morgantown, West Virginia.
A. M. JORDAN, Professor Emeritus, University of North Carolina, Chapel Hill, North Carolina.
ALFRED S. LEWERENZ, Director Emeritus of Evaluation and Research, 4107 Holly Knoll Drive, Los Angeles, California.
A. R. MEAD, Director Emeritus of Educational Research, University of Florida, 1719 N. W. 6th Ave., Gainesville, Florida.
M. J. NELSON, Dean Emeritus, Iowa State Teachers College, 6033 West Orange Drive, Glendale, Arizona.
JESSE B. SEARS, Professor Emeritus, Stanford University, Stanford, California.
CARLETON W. WASHBURNE, Chairman Emeritus, Department of Education, Brooklyn College, Brooklyn, New York.

Staff

JEAN WILLIAM
A. SCH
Bus
CLARENCE

The Journal of Educational Research

Dedicated to the Scientific Study of Education

VOLUME 56 JANUARY 1963 NUMBER 5

It Says in The JER ---- 226

A Survey of Errors in English Compositions—*Ruth Cates Baird* ---- 228

Self-Family Perceptions Related to Self-Acceptance, Manifest Anxiety, and Neuroticism—*James V. Mitchell, Jr.* ---- 236

Styles of Scholarly Production—*Robert H. Knapp, Linda A. Ehlinger* ---- 243

The Relation of Question Difficulty to Question Form—*Richard H. Bloomer* ---- 250

The Academic Achievement of Academically Talented Students —*Frederick R. Smith* ---- 255

The Image of a Teachers College as Seen By High-School Counselors—*Sidney J. Drumheller* ---- 260

Evaluation Based Upon Observation of Reports of Supervisors of Student Teaching—*Rachel D. Wilkinson* ---- 264

The Prediction of Success in Initial College Mathematics Courses—*Loren V. Corotto* ---- 268

Evaluation of a College Study Habits Course Using Scores on a Q-Sort Test as the Criterion—*James W. Creaser* ---- 272

A Method for Identifying Prospective Counselees in College —*Robert M. Roth* ---- 275

Field News—*John Schmid* ---- 277

Among Those Present ---- 279

The Journal of Educational Research is published at Madison, Wisconsin, monthly 10 times a year (the May-June and July-August issues being combined) by Dembar Publications, Box 1605, Madison 3. Wisconsin. The subscription price is $7.50 a year plus $1 a year postage for all subscriptions outside the continental United States. Single copies, $1. Address all communications to Dembar. Second class postage paid at Madison, Wisconsin.

COPYRIGHT 1963 BY DEMBAR PUBLICATIONS, MADISON 3, WISCONSIN

It Says in The JER...

MERELY GROUPING superior students in a "special" program will not result automatically in significant academic achievement, Smith says, in the absence of a carefully planned and conducted regimen designed to challenge and stimulate. He reports that the achievement of superior students grouped in so-called "enriched" courses is no greater than that of similar students in regular college-preparatory courses in the same subjects. However, students enrolled in accelerated courses achieve significantly higher STEP scores than their control groups.

GRADES ALONE are not an adequate criterion in evaluating a college-study habits course, Creaser says, in suggesting the efficacy of a Q-sort test. While improved grades may likely be the goal of some students seeking study-habits aid, they are not necessarily the goal of all, so multiple criteria must be used in any assessment.

WEAKNESSES IN STUDENT TEACHERS seem to stem from these inadequacies Wilkinson reports: poor planning and preparation, inept use of teaching materials, lack of motivational techniques, no sense of timing, unskilled questioning, untidy chores, poor class leadership, no rapport with individual children.

A SURVEY OF ERRORS in English themes suggests to Baird that (a) review of grammatical principles has little value in improving student writing, (b) an increase in the number of compositions does not necessarily produce better writing, and (c) spelling lists on a class basis are inefficient. "A student cannot write clearly without first thinking clearly," is the conclusion.

FROM AN ANALYSIS of factorial relations in 25 fields of university scholarship, Knapp and Ehlinger conclude that the traditional trilogy of “sciences,” “social sciences,” and “humanities” is a defective concept. Among the sciences, a major distinction must be made between physical and biological sciences. The homogeneity of the social sciences is largely mythical. The humanities in fact relate to sociology and political science, and do not include in their ranks arts and music. Education constitutes a special area apart from the traditional fields of scholarship.

IN PREDICTING SUCCESS in initial college mathematics courses, a general scholastic aptitude test like the ACE is acceptable if the initial course is remedial in nature, Corotto reports, but a new Mathematics Screening Test developed at the University of Houston is more diagnostic when the initial course is college-level algebra.

A STUDY OF DOCTORAL PRODUCTION in U.S. universities suggests there are four principal types of American institutions, Knapp and Ehlinger report: the Ivy League pattern, the land grant type, the metropolitan-teacher complex, and the technological. These categories are more meaningful than conventional methods of classification, the authors suggest, in “fingerprinting” the intellectual climate of an institution.

INVESTIGATING RELATIONSHIPS that may exist between perceptions of oneself within the family setting and personality variables, Mitchell concludes that self-acceptance is influenced appreciably by a satisfying and congenial family life, while the development of anxiety and neurotic symptoms will more likely be occasioned by the student’s perception that she is failing to meet standards rather than by any dissatisfaction with family life as such.

BLOOMER PROPOSES A FORMULA for determining the difficulty level of a question, independent of the material covered, which has applicability in varying question difficulty among grade levels and for children of different IQ levels.

HIGH SCHOOL COUNSELORS are more than 20 times more likely to encourage a student with high-status characteristics to attend a liberal arts college than to attend a teachers college, Drumheller reports. He says the typical counselor’s image of a teachers college is “five to ten years out of date.”

STUDENTS WHO SEEK COUNSELING from a college guidance center are characterized by a combination of poor adjustment and high anxiety, Roth reports. He suggests that prospective counselees can be identified through scores on the California Test of Personality and the Manifest Anxiety Scale.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 5, January 1963)

A Survey of Errors in English Compositions

RUTH CATES BAIRD

Oak Ridge, Tennessee, Public Schools

WHAT KINDS of errors do eleventh-grade students make in writing? Do the students make fewer of these errors after study and as the year progresses? What are the implications of the answers to these questions for the teacher of English?

This study is based on the writing of one eleventh-grade English class during 1958-59 (referred to in this study as Class A) and a comparable class during 1959-60 (Class B). These classes were similar in size: Class A, 31 students; Class B, 28 students. In the school's grouping system, both classes were "advanced"; that is, both were composed of students with intelligence quotients of 115 or above who ranked in the upper quartile on standardized achievement tests, had English grades not lower than B, and who in the opinion of previous English teachers had sufficient motivation to be successful in such a group.

Class A wrote 21 papers; Class B, 25. Thirteen of these two groups of papers on identical or similar subjects were analyzed for all errors except misspelling and confusion of terms, for which all papers were used.

Teaching methods were similar. Errors in all papers were marked and identified; papers were returned to students for inspection; errors were discussed in class. Class B kept individual records of errors made.

After the intensive study of a grammatical principle, placed at the same interval in both classes, students analyzed all previous themes for mistakes in the particular area. Sentences were corrected with the assistance of the grammar handbook.

The intensive study of punctuation followed composition number 5 (Table I). On the test following this study, grades were as follows:

Class A	Class B
A - 58%	A - 32%
B - 26%	B - 39%
C - 10%	C - 7%
D - 3%	D - 11%
F - 3%	F - 11%

However, before the study, 90 percent of Class A students made punctuation errors in themes; after the study 100 percent of these students made such errors. In Class B, 100 percent of the students made punctuation errors before the study, and the same percentage made such errors afterwards.

The study of verbs was made prior to theme number 8 (Table II). On the test given at the end of this study, grades were distributed as follows:

Class A	Class B
A - 70%	A - 41%
B - 23%	B - 37%
C - 6%	C - 19%
	D - 3%

Yet 68 percent of Class A students made verb errors before the study, and 74 percent made such errors afterwards. Following the pattern of their predecessors, 86 percent of Class B students made errors before the study, and 93 percent made verb errors afterwards.

The study of pronouns is the only one with a favorable percentage in both classes following the study between compositions numbers 9 and 10 (Table III). Grades on the pronoun test were as follows:

Class A	Class B
A - 52%	A - 57%
B - 29%	B - 32%
C - 13%	C - 7%
D - 6%	F - 4%

Before the study, 81 percent of Class A students made errors in pronoun usage; 71 percent made such errors later. Before the study, 93 percent of Class B students made such errors, and afterwards 79 percent made errors in this area. Even with improvement, the percentage of errors was still high.

TABLE I

PERCENTAGE OF STUDENTS MAKING ERRORS IN PUNCTUATION
PART 1

Theme	1		2		3		4		5		6		7		8		9		10		11		12		13	
Class	A	B	A	B	A	B	A	B	A	B	A	B	A	B	A	B	A	B	A	B	A	B	A	B	A	B
Compound sentence, coordinate conjunction	16	25	16	11	10	12	10	7	13	27		4	6	8	16	17	13	9	6	22	3	41	13	22	3	24
Introductory adverb clause	3	21		7	19	54		11	6	8	3	8		4	10	26	3	13	6	4	3	27	3		3	16
Comma before as				7									3					4					3	6		8
for						4									3	4				4			3			4
since					3																3			6		4
Quotation marks				7						8		4		4										6		16
comma								7			6	4	3	8												
Nonessential adjective clause	6														3		6	9	3	9		5				
Nonessential participial phrase	3			4																				6		
Introductory participial phrase								4																		4
Appositive								7		4	6	12			3	4	6	9	3	9	6	5				8
Parenthetic expression	6	7	6	7		8	10	11	13				3	4	10	17	3	9	3	13	6	5	10	11		8
Address		4													3	4	3			4				6	3	
Series																4						5				
Unnecessary comma	3	18	3	22	6	15	13	4	3	8	3	23		8	3	26	23	9	3	4		18	3	6		8

TABLE I
PART II

Theme	1		2		3		4		5		6		7		8		9		10		11		12		13	
Class	A	B	A	B	A	B	A	B	A	B	A	B	A	B	A	B	A	B	A	B	A	B	A	B	A	B
Colon		4						4	3		3		3			17	3			9		9				
Title book			3				32	22		4					13	13		9	6	4	23	5	35	22	13	24
story							39	7					3	4				4	19	4		27				
poem									23	4	3	8														
essay							3	4			3	12		4				4								
newspaper								4							3									6		
Comma splice		11		15	3	8			13	4		8			6	13	13	4	3	17	6	14	6		10	12
Run-on sentence	3		3	4	3	4	10		6	4				4	3	13	3		3						3	
Period, declarative sentence	3		3		3	4	3	4	3	8						9		4	3	4						4
Question mark		4		11																				6		4
Exclamation point	3																					5				

TABLE II

PERCENTAGE OF STUDENTS MAKING VERB ERRORS

Theme	1		2		3		4		5		6		7		8		9		10		11		12		13	
Class	A	B	A	B	A	B	A	B	A	B	A	B	A	B	A	B	A	B	A	B	A	B	A	B	A	B
Agreement		11	10	15	6	8		7	19	12	3	15	3	4		17	19	17	6	9	6	18	6	6		8
Tense	3	4						7	3	4	3			4	3	4	6				3			6		
Tense shift	3	18			3		3			4		4	3		10	4	3	9		9	10	18	16	11	3	4
Past participle	3	4				4								4	3			4	3	4	13		3		6	4
Split infinitive					3	8	3			4		4		4	3	9		17	3	4	3	9		6		

TABLE III

PERCENTAGE OF STUDENTS MAKING PRONOUN ERRORS

Theme	1		2		3		4		5		6		7		8		9		10		11		12		13	
Class	A	B	A	B	A	B	A	B	A	B	A	B	A	B	A	B	A	B	A	B	A	B	A	B	A	B
Agreement	3	4	3		3	8			3	8					3	17	6	22	3	13	3		6	6		4
Case	6	4				4	6	11							10	9		9					6	11	3	12
Shift of person	3	11	3	4	6	19	3	4	3	8		4						4		4			3	6		
Ambiguous reference				7	3	12	3	26		19		4		4		13		17		17		27	3	11		28
No antecedent			3				6			8	3					4	6		3	9	3	18	6		3	8
Compound personal pronoun	6	4																								
You're -- your	3		3								3															
It's -- its			3			4						4											3			4
Who -- which							3										3									

TABLE IV

PERCENTAGE OF STUDENTS MAKING ERRORS IN STRUCTURE AND ORGANIZATION

Theme	1		2		3		4		5		6		7		8		9		10		11		12		13	
Class	A	B	A	B	A	B	A	B	A	B	A	B	A	B	A	B	A	B	A	B	A	B	A	B	A	B
Sentence fragment	6	7		4		4	13	4	6	4	3	4		4	13	17	3	13	10	13	10	14	13	11		4
Faulty sentence structure	10	7	6	7	3	19	3	15	10	12		4		8	26	13	6	17	3	22	10	32	10	17	13	32
Double negative	3														3		6			4				6		
Modifiers -- dangling		7	3		3	4									13		10		3				3	11		
misplaced	3	7		4		4		4							3					13		5		6	3	4
adjective - adverb		18		7	3	8	3	4						4								9		6	3	4
Poor organization					6	4	3		16	19		8	10	13	3	9	16	13	3	22		9	6	6		4
Lack of coherence				4					3													5				
Need for transitional sentence	3								6				3											11		
Weak opening sentence					3	4		4		27		4				9		4						6		
Abrupt conclusion			3							4					3	9	26				6		3		3	
Incorrect paragraphing							3	4		4		4		4		4		4		4		5				4
Lack of unity	6	14	6		10	4			3	12				4		4	3	13	3							8
Faulty logic	3	4	10		3																	5				
Lack of clarity	6	25	10	4	16	35	23	23	39	50	10	23	6	4	19	57	29	90	3	30	3	50	29	28	10	40
Error in emphasis	3																									
Error in fact	3	4					13	4	10	31	3	4					3	17		4	6	5			3	
Error in interpretation								56	29	54	6	46	3	17	6	17		48	3	26	3	32		61		72
Omission of main idea							35	33	48	88	3	42	58	79	19	13	26	30	42	17	32	32	35	39	32	44
Omission of details, examples, illustrations	35	14	6	18	3						29	69	6	88	55	65	42	74	55	43	19	27	48	33	48	68

TABLE V

PERCENTAGE OF STUDENTS MAKING ERRORS IN WORD CHOICE

Theme	1		2		3		4		5		6		7		8		9		10		11		12		13	
Class	A	B	A	B	A	B	A	B	A	B	A	B	A	B	A	B	A	B	A	B	A	B	A	B	A	B
Inadequate vocabulary	3	18	3			8	13	18		12	6	8		8		17	13	52	3	26		14	16	17	3	16
Wordiness	3	11			3		3	7		8								13	6	9			3		13	4
Unnecessary repetition		11		4		12	3	7	3	4					3	17		4	3	13		14		5	3	
Colloquialism	6	11	3	4		12		4		4		4			13	13	3		6	9	10		10		6	8
Vulgarism							3									4					3		3			
Slang	3	4		7		4									3											4
Incorrect use of abbreviations	6	4	3		3	8	3		3							9	10	9	3					5		8
Error in writing numbers	6	21		11	3		3	4	3	8					3		10							5		4
Incorrect use of &	3					4	3	4	13									4			3					4
Incorrect use of etc.	3				3	4							3	8		4		4					3			4
Error in capitalization	13			4	10	4	6	18	3	8		4		4	10	30	23	35		22	3	14		17	3	16
Apostrophe --																										
with possessive nouns			3	7	3	4	16	7	19	27	3	4	3		16	13	6	9	10	9	3	9	3	17	3	12
unnecessary with plurals	3		3								3		3	4	3	9		4					3	5		4
unnecessary with singular verbs									3																	
necessary with contractions			3																		3	5				4
Failure to use possessive case before gerund																	3									
Use of adjective form for noun							3																			

TABLE VI

TABULATION OF TERMS CONFUSED
Figures represent numbers of students and errors

Terms Confused	Class A		Class B	
	Students	Errors	Students	Errors
1. accept - except	2	4	2	3
2. affect - effect	5	5	4	4
3. anecdotes - antidotes			1	1
4. bell - belle			1	1
5. capital - capitol			1	1
6. cites - sights	2	2		
7. clothes - cloths			2	2
8. concede - conceit			1	1
9. coarse - course	2	3		
10. expense - suspense			1	1
11. fair - fare			1	1
12. feeling - filling	1	1		
13. forth - fourth	1	1	3	4
14. hear - here	1	1		
15. human - humane			1	1
16. idea - ideal			1	4
17. knew - new	1	1	1	1
18. let's - lets	1	2	1	1
19. local - locale			2	3
20. loose - lose	3	3	1	1
21. moral - morale			1	1
22. of - off	3	4	1	1
23. pastor - pasture			1	1
24. principal - principle	2	2	1	1
25. quiet - quite	5	5	3	3
26. right - write	2	2		
27. role - roll	1	1	1	1
28. scene - seen	1	1		
29. seep - steep			1	1
30. sole - soul	2	2	1	1
31. steal - steel			1	1
32. tack - tact			1	1
33. than - then	2	7	1	4
34. pacific - specific			1	1
35. their - there	7	13	10	12
36. there - they're	1	1		
37. threw - through			2	2
38. to - too	8	15	8	13
39. too - two			2	2
40. vain - vein	1	1		
41. verses - versus			1	1
42. want - won't			1	1
43. wreak - wreck			1	1

Little consistent improvement was shown in qualities of good writing not studied periodically, but emphasized as themes were returned (Table V). This similarity to results when an intensive study was made seems to indicate that neither method is effective.

It is significant that 87 percent of Class A and 100 percent of Class B students failed to state ideas clearly (Table IV). Fuzzy thinking accounts for most of the errors in clarity.

It is important to note that 100 percent of students of both classes failed to document generalizations adequately with details, illustrations, or examples (Table IV).

Spelling errors and terms confused were listed from all papers of both classes. Terms were confused by 94 percent of Class A students and by 89 percent of Class B students (Table VI). The tabulation of these confusions is as follows:

Class A

	1 student(s)
50% of terms confused by	2
27%	3
7%	5
7%	7
3%	8
3%	14
3%	

Class B

	1 student(s)
71% of terms confused by	2
14%	3
6%	4
3%	8
3%	10
3%	

Spelling errors were made by 97 percent of Class A students and by 100 percent of Class B students. The tabulation of words misspelled is as follows:

Class A

	1 student(s)
87 % of words misspelled by	2
10 %	3
2 %	4
1/2%	5
1/4%	

Class B

	1 student(s)
84 % of words misspelled by	2
12 %	4
1+%	5
4/5%	6
1/5%	

In view of the large number of confusions and misspellings occurring only once, there are three possible conclusions: the student learned to spell or to use correctly the word upon once having it called to his attention; the student had no further occasion to use the word; or the student knew how to spell the word all along and misspelled or misused it through carelessness and failure to proofread thoroughly.

A few implications are obvious for the teacher of English:

1. A study of grammatical principles may be most successful in terms of isolated exercises, but such a study has little, if any, value in improving student writing.
2. The number of compositions written has no effect on the quality of writing. An increase in number of compositions per se does not produce better writing.
3. Spelling lists on a class basis would be a most inefficient method of helping students spell correctly.
4. Having students keep a written record of type and number of errors made has no appreciable effect on the quality of student writing.
5. A student cannot write clearly without first thinking clearly.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 5, January 1963)

Self-Family Perceptions Related to Self-Acceptance, Manifest Anxiety, and Neuroticism

JAMES V. MITCHELL, Jr.
University of Rochester

IT HAS LONG been regarded as axiomatic that the experiences of children within the family setting have a profound effect on the development of personality. But despite the general acceptance of this causal nexus, there is a paucity of significant empirical evidence relating different kinds of experiences within the matrix of family interrelationships to the development of specific personality characteristics. There are studies comparing the child-rearing practices of different cultural groups (5, 15), there are articles speculating on the probable impact of family on personality and suggesting methods of studying such relationships (10, 16), and there are some empirical studies bearing specifically on relationships between home influences and personality characteristics (1, 2, 4, 8, 11). But the latter, though increasing, are still relatively infrequent in view of the obvious importance that should be attached to the systematic mapping-out of important relationships in this area.

This critical lack of useful information about the effect of home influences on personality development has prompted G. R. Hawkes (7) to state the issue rather boldly:

> Do we really have a body of knowledge about children in families growing out of research? The answer is mostly 'No'. . . The reason for this lack is not just the relationship of psychology and sociology but a lack of methodology and theory with which to bolster the interest which many of us have.

In a later statement, Hawkes suggests the direction that future research in the area should take:

> The crux of parent-child relations as far as the child in the family is concerned appears to be in the area of children's perception of what. . . the parents are, rather than in very definite and specific characteristics of home life.

The present study attempts to use this phenomenological point of view as a framework for analyzing relationships between perceptions of family life and some important personality characteristics. Specifically, we wish to discover any relationships existing between perceptions of one's self within the family setting and the personality variables of self-acceptance, manifest anxiety, and neuroticism.

Method

The subjects for this study were 145 female students, mostly sophomores, who were taking the elementary educational psychology course as a part of the teacher preparation sequence in a midwestern university. The following instruments were administered to each of these students:

1. The Bills' Index of Adjustment and Values (3). This index consists of 49 adjectives and requires the S to rate himself in three different ways for each adjective: a) a self-description, i.e., how characteristic is the trait for this S; b) self-acceptance, i.e., how the S "feels about" himself as described for each adjective in the self-descriptive operation; c) an "ego-ideal," i.e., how much of the time the S would like to have more of the trait in question. Only the "self-acceptance" scores from the Index were used in the present study. These scores were obtained by totalling the ratings the student made on a five-point scale for each adjective: I . . . 1) very much dislike. . . . 2) dislike. . . . 3)neither dislike nor like. . . . 4) like. . . . 5) like very much. . . . being as I am in this respect.

2. A personality inventory of the "True-False" variety which included the items of the Taylor "Manifest Anxiety" scale (12) and the items from the Winne scale of neuroticism (17). Manifest anxiety and neuroticism scores were thus obtained for each subject in the study, and these scores, along with the self-acceptance score of the Bills' Index, constituted the personality variables that were related to the perceptions of self and family that were assessed by means of the instrument to be described below.

3. A fourteen-item multiple-choice inventory on perceptions of self, family, and self-within-family. The items developed for this inventory not

only represented aspects of family life and family interrelationships, but also covered aspects of the subject's general motivational orientation, since it was predicted that the variables of self-acceptance, manifest anxiety, and neuroticism would relate differently to each of these major areas in ways which could be specified by hypothesis.

The alternatives for each item in this multiple-choice inventory were ordered so as to define choices along a continuum representing greater or lesser degrees of the variable in question. An individual's "score" on a given item was simply the number assigned to the alternative chosen. Since the larger score values were associated with the positive end of the continuum for some items and the negative end for others, direction of scoring was later made uniform by subtracting certain scores from a constant. Abbreviated descriptions of the items are given in Table I, where each item is presented along with the first and last multiple-choice alternatives that were the extremes of the continuum for that item. (Item 1 is given in its entirety to exemplify the general form of each item.) Reliability data on the inventory are presented in a later section of this report.

The design of this investigation was structured in terms of certain hypotheses that were suggested by other research studies and by the author's own previous work with data of this kind. The general proposition underlying all these hypotheses was that self-acceptance is more likely to be the highest correlate of a satisfying home life, while evidence of manifest anxiety or neuroticism is more likely to be associated with attitudes toward achievement. Within this framework the following hypotheses were proposed:

1. A factor analysis of the fourteen-item multiple-choice inventory will yield two basic factors: a factor defining satisfaction with home and family, and a factor defining desire to achieve.

It was further hypothesized that if the variables of self-acceptance, manifest anxiety, and neuroticism were then separately projected on these factorial axes, the following conditions would be found to prevail:

2. Self-acceptance will be more closely related to the factor defining satisfaction with home and family than the factor defining desire to achieve.

3. Manifest anxiety and neuroticism will be more closely related to the factor defining desire to achieve than the factor defining satisfaction with home and family.

4. The factor defining satisfaction with home and family will have a higher loading for self-acceptance than for manifest anxiety and neuroticism.

5. The factor defining desire to achieve will have higher loadings for manifest anxiety and neuroticism than for self-acceptance.

In accordance with these hypotheses, the fourteen items were intercorrelated and factor-analyzed by means of the Thurstone centroid method (14). Graphical rotations were performed until an oblique simple structure was secured. The variables of self-acceptance, manifest anxiety, and neuroticism were then separately projected on the factorial axes by a method described by Fruchter (6).

Results

In conducting the factor analysis, the always difficult problem of determining how many factors to extract was solved by applying two empirical criteria to the data: Humphrey's rule (6), and Burt's empirical formula (13). Both of these indicated that a third factor should be extracted, but not a fourth. When the original centroid matrix was rotated to oblique simple structure, there resulted the factor matrix shown in Table II.

It can be observed that the first factor in this matrix has high loadings on items 1, 2, 4, 5, and 13, with lesser loadings on items 10 and 11. All of these items seem to refer to different aspects of a happy home situation—psychological "closeness," positive evaluation of parents by child, positive evaluation of child by parents, agreement on values, positive attitude toward family life in general—and it seems most logical to refer to this factorial dimension as a general "Satisfaction with Family Life" factor. The item with the highest loading on this factor (Item 2) tends to confirm this interpretation.

The second factor in the matrix has high loadings on items 6, 7, 8, and 14, with lesser loadings on items 3 and 9. Of these items, there are three that have high loadings on this factor but insignificant loadings on other factors, and these items—7, 8, and 14—refer to S's willingness to work hard on schoolwork, her rating of her study habits, and the importance she attaches to grades. This second factor is apparently best interpreted as "Academic Motivation."

The third factor has high loadings on items 3, 6, 10, and 11, with lesser loadings on items 9 and 12. It will be noted that items 3, 6, and 10 refer in turn to living up to parental expectations, parental satisfaction with S's grades, and parental satisfaction with S as a person. It is clear that the meeting of standards is common to all of these items, and it is even more clear that the standards are parent-sponsored and parent-enforced. It is, therefore, reasonable to refer to this factor as "Meeting Parental Standards."

It is evident from this analysis that the first hypothesis has been confirmed. But in addition to the two factors that were originally hypothesized, it was found necessary to extract a third factor related to the meeting of parental standards. In order to establish the reliabilities of these factors, all items having loadings of .40 or higher on a given factor were pooled together to form a measure of that factor, and corrected split-half reliability coefficients were subsequently computed for each factor. The resulting reliability coefficients were .87, .80, and .77 for the first, second, and third factors, respectively. When the relationships between these three factors are represented graphically, it can be observed that the first and second factors and the first and third factors are clearly demarcated in the common factor space, but the delimitations of the second and third factors are less clearly distinguished. The correlations between these primary factors were .28 for factors A and B, -.27 for factors A and C, and -.23 for factors B and C.

When the variables of manifest anxiety, neuroticism, and self-acceptance were projected on the original factorial axes, the loadings shown in Table III resulted. At first glance it is evident that all three variables have negligible loadings for the factor of academic motivation, but appreciable loadings for the first and third factors. It is also evident that there are some differences in factor loadings in the directions suggested by the original hypotheses. To ascertain the significance of these differences, three different tests were applied: 1) tests of the significance of differences between correlations were applied to the data of Table III, using the original factor loadings and the correlations between factors as terms in a formula described by McNemar (9); 2) for each factor the scores for the multiple-choice items having loadings of .40 or higher were added together, and these total scores were correlated in turn with manifest anxiety, self-acceptance, and neuroticism. The resulting correlations were very similar to those of Table III. The significance of differences between these correlations was then determined by the same formula applied to the earlier data; 3) fourfold tables were constructed in which multiple-choice items were classified as to whether or not they had loadings of .40 or higher on a particular factor, and whether they were significantly related to a given personality variable as determined by analysis of variance technique. Chi-square values were then computed to determine any relationship that might exist between an item's loading on a particular factor and its relationship to one of the personality variables. These several tests yielded the following results:

1. The relationship between self-acceptance and the factor of satisfaction with family life is significantly greater than that between manifest anxiety and the same factor. This is true for all three tests applied. This same relationship between self-acceptance and the factor of satisfaction with family life is also significantly greater than that between neuroticism and the same factor, and this is again confirmed by all three tests. It thus appears that the fourth hypothesis has ample confirmation.

2. The relationship between manifest anxiety and the factor of meeting parental standards appears to be significantly greater than that between manifest anxiety and the factor of satisfaction with family life. Although this difference is not quite significant for the second test applied, it is significant for the first and third. Neuroticism again shows a close resemblance to manifest anxiety, for the relationship between neuroticism and the factor of meeting parental standards is significantly greater than that between neuroticism and the factor of satisfaction with family life, and this is true for all three tests applied. Thus it appears that whether we regard the third hypothesis as confirmed or not depends largely on whether we can regard the third factor extracted as the "desire to achieve" factor that is specified in this third hypothesis. Since both the second and third factors are achievement-related factors, the issue is subject to conjecture. But there can be little doubt that manifest anxiety and neuroticism have significantly higher correlations with the factor of meeting parental standards than with the factor of satisfaction with family life.

3. The confirmation of the second hypothesis again depends upon our interpretation of the "desire to achieve" factor specified in the hypothesis. If the second factor is so regarded, then the loading of self-acceptance on the first factor is significantly higher than the loading of self-acceptance on the second factor, and the hypothesis is confirmed. But if the third factor is so regarded, the second hypothesis is unconfirmed, for although the correlation of self-acceptance with the first factor is slightly higher than that between self-acceptance and the third factor, this difference is not significant. It should be noted, however, that the data still permit us to say that self-acceptance is at least if not more closely related to the factor of satisfaction with family life than the factor of meeting parental standards.

4. Although there are a few differences in the specified direction, the fifth hypothesis received no statistical confirmation.

Discussion

In interpreting the data of this study there are certain results that are worthy of further comment and elaboration:

1. The fact that two factors related to motivation were extracted from the fourteen-variable correlation matrix rather than just the one expected suggests that variables like "achievement motivation" and "desire to achieve" are much more complex than is commonly realized. In this study, for instance, it was apparent that the S's perception of how hard she actually worked was factorially distinguishable from her perception of whether or not she was meeting parental standards. Furthermore, the factor of meeting parental standards had substantial loadings for the variables of self-acceptance, manifest anxiety, and neuroticism, while the factor of academic motivation had negligible and insignificant loadings for these variables. This suggests that the emotional and adjustmental correlates

of an individual's motivational orientation are more likely to stem from the S's perception of how authority figures evaluate herself and her work than from her perception of how much effort she actually exerts in her work. The latter is probably more closely related to actual accomplishment, but the former is undoubtedly more reflective of the psychological meaning of an individual's characteristic achievement orientation within her total adjustment pattern.

2. A most important set of findings resulted when the variables of manifest anxiety, neuroticism, and self-acceptance were projected on the original factor axes. Here it became evident that the correlates of self-acceptance on the one hand and manifest anxiety and neuroticism on the other were quite different in nature. The quality and atmosphere of home life is more closely related to self-acceptance than it is to manifest anxiety or neuroticism. And manifest anxiety and neuroticism are more related to the S's perception of whether she has met parental standards than it is to the interpersonal congeniality of her home life. Looking at it in a slightly different way, we can say that self-acceptance is not only the highest correlate of a happy home condition, but is also as closely or more closely related to congenial family living than it is to the meeting of parental standards. The meeting of parental standards, on the other hand, is equally if not more highly correlated with manifest anxiety than it is with self-acceptance. These findings suggest the general conclusion that

> self-acceptance is influenced appreciably by a satisfying and congenial family life, while the development of anxiety and neurotic symptoms will more likely be occasioned by the subject's perception that she is failing to meet standards rather than any dissatisfaction with family life per se.

Doubtlessly such interpretations will have to be considered with the usual caution about direction of causality. It might be argued, for instance, that the loading of self-acceptance on the first factor was not due to the benign influence of a satisfying family life on the development of self-acceptance but rather due to the fact that a self-accepting person is more inclined to be accepting of all features of his environment. If this were the case, however, it would be difficult to explain why this generally acceptant attitude did not result in a high loading for the factor of academic motivation as well. Similar counter-arguments could be offered against other alternative explanations.

3. Although it may be of lesser importance than the foregoing data, certain interesting tendencies emerge when the projected variables of Table III are considered in relation to the correlations between primary axes that were reported earlier. The factor of satisfaction with family life is not only related to self-acceptance, but the same factor is also positively related (.28) to the factor of academic motivation and negatively related (-.27) to the factor of meeting parental standards. This suggests that a satisfying family life may not only have a facilitating influence on self-acceptance but may also be instrumental in encouraging fairly rigorous standards which the individual may not feel that she has always attained but for which she has the confidence and motivation to try to attain. The interpretation receives partial confirmation from the correlations between the original multiple-choice items, where item 1 correlates negatively with items 3 and 6 but has a significantly positive correlation with item 10.

Summary

The purpose of this study was to investigate relationships that might exist between perceptions of oneself within the family setting and the personality variables of self-acceptance, manifest anxiety, and neuroticism. One hundred forty-five female students in educational psychology were administered: a) a fourteen item multiple-choice inventory on perceptions of self, family, and self-within-family; b) the Bills' Index of Adjustment and Values, from which "self-acceptance" scores were obtained; and c) a personality inventory of the true-false variety which included the items of the Taylor Manifest Anxiety Scale and the Winne scale of neuroticism. The fourteen-item multiple-choice questionnaire was factor analyzed, and the variables of self-acceptance, manifest anxiety, and neuroticism were separately projected on the factorial axes.

The factor analysis yielded three factors, identified as A) Satisfaction with Family Life, B) Academic Motivation, and C) Meeting Parental Standards. The latter two were found to be factorially distinguishable despite their close relationship as achievement-oriented variables. The projected personality variables had appreciable loadings for factors A and C but negligible or zero loadings for factor B. Self-acceptance had a significantly higher loading on Factor A than manifest anxiety or neuroticism. The loading of manifest anxiety on factor C was significantly higher than its loading on factor A. This was also true for neuroticism. These and allied results were interpreted to mean that self-acceptance is influenced appreciably by a satisfying and congenial family life, while the development of anxiety and neurotic symptoms will more likely be occasioned by the S's perception that she is failing to meet standards rather than any dissatisfaction with family life per se. Consideration of the projected personality variables along with the correlations between the factorial axes suggested that a satisfying family life may not only facilitate self-acceptance but may also increase motivation for schoolwork.

REFERENCES

1. Baldwin, A. L. "The Effects of Home Environment on Nursery School Behavior," Child Development, XX (1949), pp. 49-61.
2. Beals, L. "A Study of Certain Home Factors and Their Relationship to the Personal Adjustment of Children," School and Society, LXXII (1950), pp. 55-57.
3. Bills, R. E. "An Index of Adjustment and Values," Journal of Consulting Psychology, XV (1951), pp. 257-61.
4. Carter, D. C. "The Influence of Family Relations and Family Experiences on Personality," Marriage and Family Living, XVI (1954), pp. 212-15.
5. Davis, A., and Havighurst, R. J. "Social Class and Color Differences in Child Rearing," American Sociological Review, XI (1946), pp. 698-710.
6. Fruchter, B. Introduction to Factor Analysis (New York: Van Nostrand, 1954), appendix and pp. 79-80.
7. Hawkes, G. R. "The Child in the Family," Marriage and Family Living, XIX (1957), pp. 46-51.
8. Jourard, S. M., and Remy, R. M. "Perceived Parental Attitudes, the Self, and Security," Journal of Consulting Psychology, XIX (1955), pp. 364-66.
9. McNemar, Q. Psychological Statistics (New York: John Wiley and Sons, 1949), pp. 255-62; 124-25.
10. Sanford, R. N. "Family Impact on Personality: The Point of View of a Psychoanalyst," in Problems in Social Psychology, J. E. Julett and R. Stagner, eds., (Urbana: University of Illinois, 1952).
11. Sears, R. R., Maccoby, Eleanor E., and Levin H. Patterns of Child Rearing (Evanston, Ill.: Row, Peterson and Co., 1957).
12. Taylor, Janet A. "A Personality Scale of Manifest Anxiety," Journal of Abnormal Psychology, XXXVIII (1953), pp. 285-90.
13. Thomson, G. The Factorial Analysis of Human Ability (Boston: Houghton-Mifflin, 1951), pp. 122-23.
14. Thurstone, L. L. Multiple Factor Analysis (Chicago: University of Chicago Press, 1947).
15. Whiting, J. W. M., and Child, I. L. Child Train-and Personality (New Haven: Yale University Press, 1953).
16. Winch, R. F. "The Impact of Family Upon Personality: the View of One Sociologist," in Problems in Social Psychology, J. E. Hulett and R. Stagner, eds., (Urbana: University of Illinois, 1952).
17. Winne, J. F. "A Scale of Neuroticism: an Adaptation of the Minnesota Multiphasic Personality Inventory," Journal of Clinical Psychology, VII (1951), pp. 117-22.

TABLE I

ABBREVIATED DESCRIPTIONS OF ITEMS IN THE MULTIPLE-CHOICE INVENTORY

Item Number	Item Description
1.	Which of the following best describes your own family during that period of time when you were growing up? 1. We were a very "close" family. 2. We were probably closer to one another than would occur in a typical or average family. 3. Our relationships were as close as those occurring in the typical or average family. 4. By and large each member of the family tended to go his separate way. 5. We were a family only in name; nobody cared much about anyone else.

TABLE I (Continued)

Item Number	Item Description
2.	Let's assume that at some time in the future you will get married and have a family. Compare the circumstances in which you grew up with the kind of home situation which you would like to provide for your own children, and then select the one statement that best represents the results of your comparison. (From "1. There will be some major differences between the home situation I would provide for my youngsters and the one I had" to "3. I would make no changes at all; I believe the conditons under which I grew up were fine in every respect.")
3	Do you feel that thus far you have really lived up to your parents' expectations for you? (From "1. Yes; I have not only lived up to their expectations; I have actually exceeded them" to "4. No; I have fallen far short of their expectations for me.")
4.	Compare your own mother with the "ideal" mother. Which of the statements below best represents the results of your comparison? (From "1. There are some major differences between my mother and the 'ideal' mother" to "3. There are no differences at all between my mother and the 'ideal' mother.")
5.	Compare your own father with the "ideal" father. Which of the statements below best represents the results of your comparison? (Alternatives similar to those for item 4.)
6.	Are your parents satisfied with your grades in college? (From "1. Extremely satisfied and proud" to "4. Not satisfied at all; they are rather distrubed about the grades I've been getting.")
7.	How hard do you work on your schoolwork, in comparison with the typical student of your sex? (From "1. I probably work a great deal harder than the typical student" to "5. I don't work nearly as hard as the typical student.")
8.	How important are grades to you? (From "1. Extremely important" to "4. Very unimportant; I have very little interest in what grades I get.")
9.	How much intellectual curiosity do you think you have, in comparison with the typical student of your sex? (From "1. I have much more intellectual curiosity than the typical student" to "5. I have much less intellectual curiosity than the typical student.")
10.	Do you feel that your parents are satisfied with you as a person? (From "1. Extremely satisfied" to "4. Very dissatisfied; they would like to have a major overhaul.")
11.	Are you satisfied with yourself as a person? (From "1. Extremely satisfied" to "4. Very dissatisfied; there are many, many things I would like to change in myself.")
12.	Did your parents encourage you to be "on your own" more than the average boy or girl? (From "1. Yes; I think such independence was encouraged much more than it is for the average child" to "5. Definitely no; my parents actually seemed to discourage me from being 'on my own.'")
13.	Do you and your parents agree on "values," i.e., what constitutes the really important things in life? (From "1. We are in complete agreement" to "4. We usually are in complete disagreement about the important values in life.")
14.	How would you rate your study habits in comparison with the average student of your sex? (From "1. Much better than average" to "4. Much poorer than average.")

TABLE II

FACTOR MATRIX FOR THE FOURTEEN ITEM MULTIPLE-CHOICE INVENTORY

Item	Factor A Satisfaction with Family Life	 B Academic Motivation	 C Meeting Parental Standards
1.	.56	-.03	-.01
2.	.66	-.05	.00
3.	-.01	.33	.61
4.	.50	.10	.11
5.	.54	-.01	-.01
6.	-.12	.58	.60
7.	-.03	.81	.11
8.	.03	.50	-.05
9.	.02	.35	.44
10.	.46	.10	.69
11.	.42	-.07	.67
12.	.19	.00	.45
13.	.55	-.07	.11
14.	.06	.71	.31

TABLE III

FACTOR LOADINGS OF SELF-ACCEPTANCE, MANIFEST ANXIETY, AND NEUROTICISM AFTER PROJECTION ON THE ORIGINAL FACTORIAL AXES

Personality Variable	Factor A Satisfaction with Family Life	 B Academic Motivation	 C Meeting Parental Standards
Self-Acceptance	.40	.01	.37
Manifest Anxiety	-.25	.12	-.44
Neuroticism	-.13	.00	-.35

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 5, January 1963)

Styles of Scholarly Production*

ROBERT H. KNAPP
LINDA A. EHLINGER
Wesleyan University

THE LAST DECADE has seen an almost obsessive concern among college administrators and professors in the classification and evaluation of American institutions of higher education. This ferment has been actuated by the desire to evaluate comparative achievements, and in part to plot new courses of institutional development. Such concerns have inevitably raised the question of types and varieties of American higher educational institutions. Thus, institutions have been variously classified. This writer has classified them as public and private, universities and colleges, church-controlled and non-denominational, and by region, cost and size, upon the a priori assumption that these represent meaningful categories (1, 2). The present article seeks in a modest way to approach the problem of types and varieties of institutions of higher education in a quite different manner, namely, in terms of the type of scholar produced by such an institution. It is, the authors believe, unique that this typology rests not upon a priori categories but rather upon statistical evidence from which an inductive classification is derived.

The data employed in the present study are taken from Doctorate Production in United States Universities 1936-1956, National Academy of Sciences (3). The authors of this publication have selected the 95 institutions yielding the greatest number of graduates who have subsequently taken a Ph.D. within the time span indicated. For each of these 95 institutions, the number of Ph.Ds. taking their baccalaureate degrees from each of 31 fields of scholarship is recorded. In the development of this study these numbers were combined under five main headings, namely, Physical Sciences, Biological Sciences, Humanities, Social Sciences, and Education. It should also be noted that the authors took the liberty of eliminating certain categories and combining others. Thus, for example, history was classified among the humanities and business and commerce were eliminated from this category. Next, for each institution, the percentage of total baccalaureates who had taken doctorates in each of the five major categories was calculated. The resulting table of percentages for each institution yielded a profile which revealed, immediately upon inspection, the areas of their strength and weakness in the production of graduates who had earned their Ph.D. It should be emphasized here that the mode of analysis ignores the absolute number, and also ignores the ratio of scholars to graduates. It deals only with relative achievement in different areas of scholarship.

In the analysis of the resulting data an inverse factor analysis in which each institution was correlated with every other, yielding a 95 × 95 correlation matrix, was performed. The correlations were based, to be sure, on only five determinations, but these, especially in the case of larger institutions, could each be determined with considerable precision. The authors did not undertake this analysis in terms of the 30-odd fields listed because the incidence of scholars in many institutions was so low that no reliability could be assigned to such percentages. The correlation matrix described above was subjected to a centroid factor analysis and subsequently the axes were rotated according to a Varimax program, yielding four orthogonal factors. Table I shows the rotated factor loadings of each of the 95 institutions on the four factors identified.

For each of the factors the authors have selected a representative number of institutions with high saturations and calculated the mean percentage of their scholars in each of the five categories. Figure 1 shows the profile of each of the four groups of selected institutions in terms of their percentage deviation from the average.

It will be seen that the first group of institutions

---

* This research was supported by the Council of the Humanities of Princeton University under a grant from the Ford Foundation.

TABLE I

FACTORS EMERGING FROM VARIMAX ROTATION ON 95 SELECTED INSTITUTIONS*

Institutions	Factor I Ivy League	Factor II Land Grant	Factor III Metropolitan-Teacher	Factor IV Technological
University of California, Berkeley	+.44	+.73	-.01	+.52
City College of New York	+.43	+.16	+.52	+.72
University of Illinois	+.36	+.54	+.09	+.76
University of Chicago	+.84	+.25	+.17	+.46
University of Wisconsin	+.31	+.93	+.05	+.20
Harvard University	+.90	-.08	-.01	+.44
University of Minnesota	+.33	+.87	+.12	+.34
Columbia University	+.03	-.48	+.69	+.53
University of Michigan	+.61	+.36	+.12	+.70
New York University	+.01	-.05	+.94	+.34
Cornell University	+.26	+.86	-.16	+.41
Ohio State University	+.25	+.78	+.28	+.50
University of California, Los Angeles	+.82	+.24	+.46	+.26
Yale University	+.84	-.13	-.01	+.54
University of Washington	+.52	+.47	+.26	+.67
Brooklyn College	+.71	+.29	+.33	+.56
Pennsylvania State University	+.17	+.79	+.12	+.58
Massachusetts Institute of Technology	+.29	+.20	+.06	+.94
University of Texas	+.75	+.12	+.19	+.63
University of Pennsylvania	+.60	+.02	+.09	+.80
Stanford University	+.79	+.14	+.44	+.41
Northwestern University	+.79	+.06	+.30	+.53
State University of Iowa	+.90	+.24	+.34	-.11
Purdue University	+.32	+.61	+.03	+.73
Oberlin College	+.67	-.15	-.02	+.73
Princeton University	+.75	-.29	-.16	+.57
Iowa State College of Agriculture and Mechanic Arts	+.13	+.87	-.11	+.47
University of Nebraska	+.10	+.99	+.12	-.06
Catholic University	+.25	-.86	-.41	-.21
University of Missouri	-.01	+.93	+.37	+.07
University of Pittsburgh	+.38	+.33	+.82	+.30
Rutgers University	-.05	+.95	-.12	+.29
Indiana University	+.45	+.32	+.57	+.61
University of Utah	+.45	+.44	+.22	+.75
Syracuse University	+.41	+.59	+.69	+.09
University of Kansas	+.54	+.59	+.21	+.56
Michigan State University	-.06	+.97	-.24	-.07
California Institute of Technology	+.28	+.23	+.03	+.93
Johns Hopkins University	+.38	+.24	+.06	+.89
Dartmouth College	+.95	-.04	-.01	+.32
University of North Carolina	+.91	-.13	-.05	+.41
University of Colorado	+.61	+.35	+.13	+.70
Utah State Agricultural College	+.03	+.96	+.01	-.30
DePauw University	+.51	+.57	+.14	+.63
University of Rochester	+.46	+.17	-.10	+.87
Brown University	+.61	+.24	-.02	+.75
Carnegie Institute of Technology	+.24	+.12	+.11	+.96
Brigham Young University	-.25	+.95	- 00	-.17
Washington University	+.76	+.23	+.11	+.60
Wayne State University	+.48	00	+.70	+.54
University of Florida	+.23	+.87	+.26	+.36
Boston University	-.15	-.52	+.73	-.41
Temple University	+.31	+.13	+.85	+.42
Swarthmore College	+.89	+.17	+.05	+.43

---

* Arranged in order of total number of scholars produced.

TABLE I (continued)

	Factor I	Factor II	Factor III	Factor IV
Institutions	Ivy League	Land Grant	Metropolitan-Teacher	Technological
University of Oklahoma	+.45	+.22	+.47	+.73
University of Southern California	+.38	+.02	+.75	-.52
University of Kentucky	+.31	+.83	+.11	+.45
Kansas State College	+.06	+.99	-.14	+.08
Hunter College of New York	-.06	-.40	+.35	-.85
University of Maryland	+.03	+.94	-.20	+.30
University of Massachusetts	-.12	+.91	-.37	-.13
Louisiana State University	+.13	+.91	00	+.41
Miami University	+.78	+.47	+.09	+.40
George Washington University	+.66	+.74	+.08	+.11
University of Virginia	+.82	-.11	-.12	+.56
Western Reserve University	+.94	+.30	+.09	+.06
University of Cincinnati	+.45	+.04	+.28	+.85
State College of Washington	-.05	+.98	-.09	+.19
Oregon State College	-.11	+.89	-.04	+.45
West Virginia University	+.10	+.79	+.09	+.60
Amherst College	+.91	-.38	-.02	+.15
Oklahoma Agricultural and Mechanical College	-.29	+.93	-.09	-.20
Fordham University	+.98	+.13	+.14	+.08
University of Notre Dame	+.43	+.08	+.03	+.90
University of Buffalo	+.33	+.27	+.37	+.83
Duke University	+.84	+.08	+.34	+.43
University of Denver	+.11	-.02	+.95	-.25
University of Alabama	+.36	-.27	+.89	+.07
Rensselaer Polytechnic Institute	+.26	+.19	+.07	+.95
Wesleyan University	+.89	-.23	-.15	+.37
Reed College	+.71	+.16	+.19	+.66
Emory University	+.61	+.14	-.27	+.74
St. Louis University	+.94	-.24	+.04	+.24
College of Wooster	+.61	+.02	+.03	+.80
Ohio University	+.24	+.27	+.86	+.36
University of New Hampshire	+.09	+.80	+.04	+.59
North Texas State College	-.56	-.04	+.79	+.16
University of Oregon	-.09	-.48	+.85	-.19
Iowa State Teachers College	-.46	-.18	+.84	-.20
University of Tennessee	+.02	+.93	+.20	+.31
University of Idaho	-.14	+.97	-.15	+.09
Agricultural and Mechanical College of Texas	+.10	+.89	-.12	+.43
University of Georgia	-.11	+.99	00	-.11
University of Connecticut	+.09	+.96	-.10	-.23
University of Arkansas	+.44	+.77	+.33	+.33

is notably high in the humanities, strong in the social sciences, moderately productive in the physical and biological sciences, but notably low in the field of education. The institutions reported in Table II had factor loadings on the first factor exceeding +.80. It will be observed that in this group Ivy League institutions heavily predominate, and that only three of the 15, namely, University of California, Los Angeles; State University of Iowa, and University of North Carolina, are publicly supported. It would seem warranted to describe this factor as "Ivy League."

The second factor shows a very heavy contribution in the field of biological sciences, a modest contribution in the field of education, and a notably low contribution in the humanities. The 23 institutions listed in Table III yielded factor loadings in excess of +.85. Inspection of this listing will immediately reveal the heavy predominance of land-grant institutions in the Middle and Far West. It will also be noted that all are publicly supported, with the exception of BrighamYoung and Cornell, the latter being a special case in that it is both publicly and privately sponsored. This factor can be designated as a "land-grant" factor.

The third factor is represented in Figure 1 by 12 institutions with factor loadings in excess of +.75, as shown in Table IV. This group of institutions is particularly weak in both physical and biological sciences and notably strong in the field of education. Their contributions to the social sciences are clearly superior and their contribution to the humanities substantial. There would appear to be two types of institution included here: first, institutions located in metropolitan centers and subsequently sponsored by such, and secondly, teachers colleges, or institutions which were earlier teachers colleges. It would seem appropriate to describe this factor as "Metropolitan-Teacher."

The fourth factor is notably represented by nine institutions showing factor loadings exceeding +.85, as shown in Table V. Among this group of institutions, their sphere of special attainment is clearly in the physical sciences, though, interestingly, their contribution to the humanities appreciably exceeds that of the land-grant institutions. Here we have a number of better-known engineering and scientific institutions such as M.I.T., California Institute of Technology, Carnegie Institute of Technology, and Rensselaer Polytechnic Institute, as well as a number of institutions of strong technological leanings such as the University of Rochester, University of Buffalo, Johns Hopkins, etc. This factor may appropriately be called a "technological" factor.

There are, then, in summary, four model styles of scholarly production identified among our 95 institutions most productive of graduates who have taken Ph.Ds. from 1936-1956. These are, respectively, an Ivy League, a land-grant, a metropolitan-teacher, and a technological style. They differ radically in the pattern of their scholarly output and constitute, we believe, a fruitful basis for evaluating the general character and intellectual climate of a particular institution. It may be profitable to examine one or two cases that are not included among those selected as most typical of the four factors. For example, the University of Pennsylvania, which is normally classified as an Ivy League institution, actually yields a higher loading on the technological factor than on the Ivy League factor because of its relatively high standing in the physical sciences. As such, it constitutes a kind of hybrid. The same general pattern characterizes the University of Colorado, though it might commonly be assumed that these two institutions are profoundly different, as indeed they are in certain respects. The University of Indiana and the University of Washington are interesting because they have no distinguishing factor loadings, showing rather a moderate loading on all factors. They constitute a sort of "All-American" type of institution. Another institution of particular interest is Catholic University, which yields a uniquely high negative loading on the land-grant factor. This suggests that its pattern of scholarly attainments is antithetical to that for the land-grant schools.

Thus, any of these 95 institutions may be evaluated with respect to the modal styles of scholarly production. Fortunately, this study is not limited to the 95 most productive institutions. The publication of the National Academy of Sciences (3) gives the record for almost all other undergraduate institutions in the nation, and it is possible, by consulting these figures, to determine the degree to which any institution approximates each of these four styles or represents a hybrid of them.

It seems to the authors that the measure directly reflects the actual climate of an institution and that it is a more useful mode of classifying institutions for many purposes than the more conventional categories earlier enumerated in this article. Though the number of scholars produced by many institutions is a relatively small proportion of their total graduates, still it is believed that their distribution across the major fields of learning serves as a kind of fingerprint through which the intellectual climate of an institution may be discerned. Classification according to the typology taken in conjunction with some measure of overall productivity, i.e., the rate of scholars to graduates, offers, it is believed, a very quick and meaningful picture of the institution as a whole.

This particular analysis was undertaken for a more specific purpose than the devising of a typology based upon scholarly output. The authors were concerned with identifying the type of institution in which the production of humanistic scholars constitutes a major area of achievement. In

TABLE II

FACTOR I — IVY LEAGUE
PERCENTAGES IN EACH OF FIVE AREAS

Institutions	I Physical Sciences	II Biological Sciences	III Social Sciences	IV Humanities	V Education
University of Chicago	34.6	16.2	29.3	11.7	8.2
Harvard University	35.5	10.3	31.1	18.7	4.4
University of California, Los Angeles	30.3	15.2	32.4	9.0	13.1
Yale University	37.6	8.6	29.2	19.5	5.1
State University of Iowa	22.9	18.4	32.7	13.0	13.0
Dartmouth College	29.1	14.4	29.1	18.8	8.6
University of North Carolina	29.4	13.5	27.3	20.2	9.6
Swarthmore College	36.3	14.8	31.4	13.6	3.9
University of Virginia	36.4	11.0	26.8	21.0	4.8
Western Reserve University	25.7	19.6	30.8	14.5	9.4
Amherst College	27.1	7.8	33.7	24.0	7.4
Fordham University	22.2	18.8	24.4	18.4	16.2
Duke University	33.1	12.4	30.6	12.4	11.6
Wesleyan University	29.1	12.6	27.4	22.6	8.3
St. Louis University	28.3	10.8	31.4	20.6	9.0
Mean	30.51	13.63	29.84	17.20	8.84

TABLE III

FACTOR II — LAND GRANT
PERCENTAGES IN EACH OF FIVE AREAS

Institutions	I Physical Sciences	II Biological Sciences	III Social Sciences	IV Humanities	V Education
University of Minnesota	29.4	28.3	21.3	7.5	13.3
University of Wisconsin	26.9	30.2	21.4	8.2	13.3
Cornell University	34.7	35.3	16.5	6.3	7.2
Iowa State College of Agriculture and Mechanic Arts	40.8	40.7	11.7	0.2	6.6
University of Nebraska	22.1	32.7	21.2	7.4	16.6
University of Missouri	22.7	27.3	20.4	9.2	20.4
Rutgers University	29.7	39.3	12.5	5.2	13.3
Michigan State University	21.6	52.6	13.9	2.9	9.0
Utah State Agricultural College	16.1	45.5	22.9	1.4	14.1
Brigham Young University	18.2	38.1	16.4	7.7	19.6
University of Florida	29.8	28.2	20.5	5.3	16.2
University of Maryland	33.7	45.8	10.9	2.2	7.4
Kansas State College	27.1	46.2	16.3	1.5	8.9
University of Massachusetts	17.6	64.6	9.4	4.1	4.4
Louisiana State University	30.6	31.6	16.8	7.1	13.8
State College of Washington	27.4	39.5	14.2	5.0	13.9
Oregon State College	36.3	40.7	7.7	0.7	14.7
Oklahoma Agricultural and Mechanical College	15.6	49.4	13.0	3.3	18.6
University of Tennessee	27.9	30.7	17.7	5.6	18.1
University of Idaho	24.9	45.0	12.0	4.3	13.9
Agricultural and Mechanical College of Texas	39.4	42.3	11.3	0.0	7.0
University of Georgia	20.1	35.8	18.1	8.3	17.6
University of Connecticut	18.1	48.6	22.4	1.4	9.5
Mean	26.55	39.93	16.02	4.56	12.93

TABLE IV

FACTOR III — METROPOLITAN-TEACHER PERCENTAGES IN EACH OF FIVE AREAS

Institutions	I Physical Sciences	II Biological Sciences	III Social Sciences	IV Humanities	V Education
New York University	25.5	12.6	23.0	10.4	28.5
University of Pittsburgh	26.5	17.1	25.8	8.4	22.2
Temple University	29.2	12.9	25.8	7.4	24.6
University of Denver	17.7	14.6	29.2	9.7	28.8
University of Alabama	21.6	14.1	25.3	15.8	23.2
Ohio University	29.3	14.9	26.0	4.2	25.6
University of Oregon	16.0	8.0	27.2	15.5	33.3
Iowa State Teachers College	9.6	8.6	23.4	4.6	53.8
Wayne State University	29.2	12.8	24.9	12.3	20.8
Boston University	10.7	8.4	28.5	17.6	34.8
University of Southern California	14.2	16.2	34.6	10.7	24.3
North Texas State College	21.2	12.9	14.7	6.0	45.2
Mean	20.89	12.76	25.70	10.22	30.42

TABLE V

FACTOR IV — TECHNOLOGICAL PERCENTAGES IN EACH OF FIVE AREAS

Institutions	I Physical Sciences	II Biological Sciences	III Social Sciences	IV Humanities	V Education
Massachusetts Institute of Technology	89.7	6.4	2.8	0.2	0.9
California Institute of Technology	86.7	8.7	2.3	1.1	0.2
Johns Hopkins University	55.1	13.9	14.7	8.7	7.6
University of Rochester	40.7	16.0	17.1	16.6	9.6
Carnegie Institute of Technology	85.5	1.4	2.6	3.1	7.4
University of Cincinnati	40.9	10.5	20.6	12.6	15.4
University of Notre Dame	54.3	9.3	15.8	13.4	7.3
University of Buffalo	41.9	14.5	19.0	7.3	17.3
Rensselaer Polytechnic Institute	91.3	5.4	0.8	0.0	2.5
Mean	65.50	9.68	10.63	7.20	7.58

this connection it has been shown that the Ivy League institutions excel in this field, confirming conclusions emerging from earlier studies (1, 2). But the scholar or administrator interested in other areas of scholarly distinction may find the examination of the data fruitful as well.

REFERENCES

1. Knapp, R. H., and Goodrich, H. B. The Origins of American Scientists (Chicago: University of Chicago Press, 1953).
2. Knapp, R. H., and Greenbaum, J. J. The Younger American Scholar (Chicago: University of Chicago Press, 1954).
3. Doctorate Production in United States Universities 1936-1956, Publication No. 582 (Washington, D.C.: National Academy of Sciences, National Research Council).

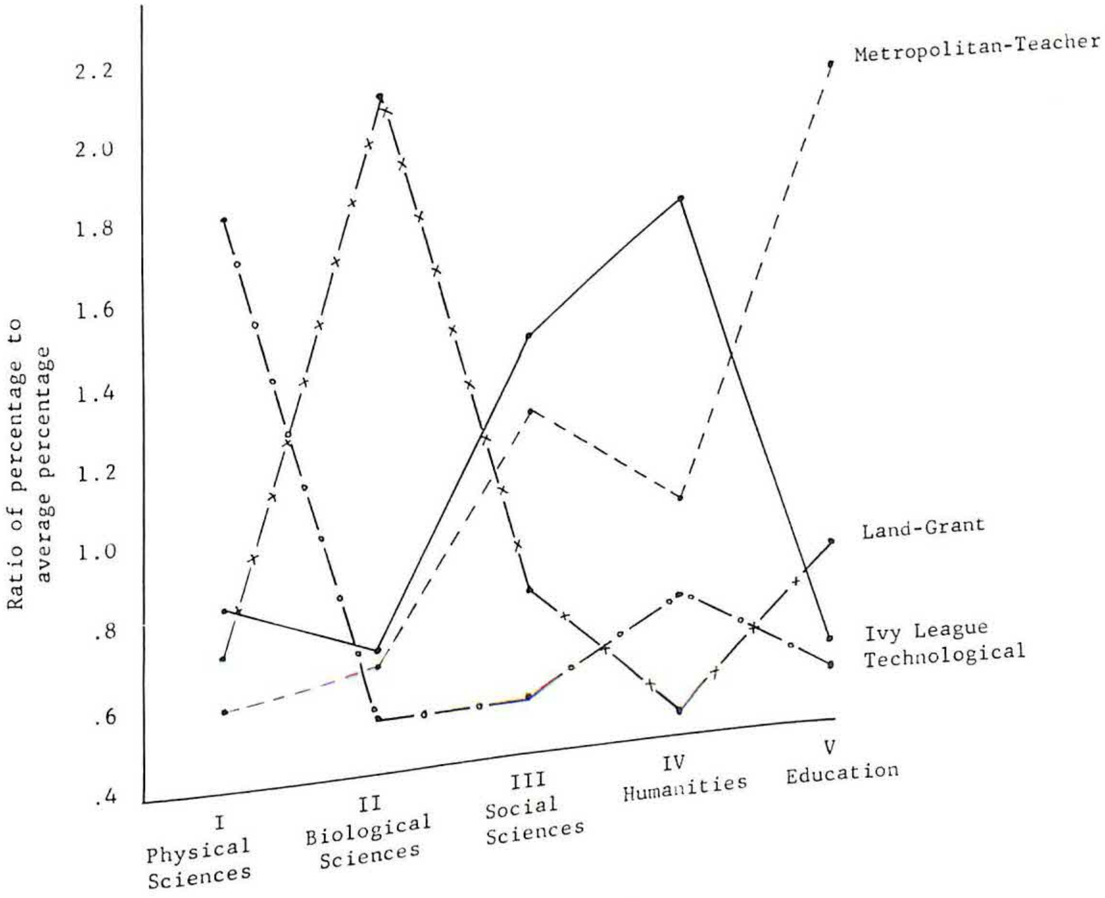


FIGURE 1

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 5, January 1963)

The Relation of Question Difficulty to Question Form

RICHARD H. BLOOMER
State University College of Education
Geneseo, New York

THE FORM OF A QUESTION varies with the grade level of the child answering the question. A number of variables in the classroom questioning situation seems to affect the difficulty of questions. The present paper examines a number of these variables to determine their relationship to question difficulty.

Literature

Competition. Sims (7) found that children respond better under conditions of competition; group rivalry is apparently more effective for producing learning than no rivalry at all. Individual rivalry is superior to group rivalry. Similarly, Hurlock (3) found that children showed consistent improvement in learning when publicly praised. Adverse effects occurred when they were reproved or ignored.

Competition, inducing greater increments of learning when success occurs, at the same time decreases the probability of success. For present purposes, then, increasing competition will be considered to increase the difficulty of a given question.

The Size of the Answer. The work of Ebbinghaus (2) evidences that an increase in the number of nonsense syllables increases the number of trials necessary to learn each syllable. Where meaningfulness was a factor, Lyons (5) found that the number of repetitions increased with the size of the learning problem, but that this increase was not proportional. In this same area, Cofer (1) found that when an answer in specific terms was required, more trials were needed than when an answer in non-specific terms was required, and further, that with a specific answer, the effects of length were greater. For present purposes, then, the longer the required answer, the more difficult the question; and length is of greater effect when the answer is specific.

Abstraction of the Question. Pfafflin (6) found that the more meaningful the question, the smaller the number of errors learned. Her stimuli were pictures with three degrees of meaningfulness which were to be matched with both relevant and irrelevant labels. Fewer errors were made for the relevant labels, indicating that relevancy is a factor in question difficulty.

The Form of the Answer. Luh (4) and numbers of others testing retention under conditions of recognition and recall found that recognition is superior. By the same token, Sisson (8) found that material already organized is easier to learn than material requiring the learner to organize as well as to learn the items. These data imply that a question requiring recognition of the answer is easier than a question requiring recall of the answer, which in turn is easier than a question requiring both recall and reorganization of the material.

The literature thus pointed to a number of variables applicable to the determination of question difficulty.

Procedure

The Criterion of Question Difficulty. In the present study, question difficulty was determined from the grade level assigned a given question by commercial workbook publishers. The 172 questions were selected at random from eight successive grade levels of reading and language arts workbooks by eight publishers. Each of these questions was assigned a serial number and all reference to its grade level was deleted. The applicability of each variable to each question was then determined by two scorers: an expert scorer who was familiar with the scoring system and had used it on other occasions, and a group of several students who were simply given the scoring instructions and told to determine a total difficulty score.

The Variables. The following scoring categories evolved from application of relationships found in the literature. Weightings for each category were determined from an examination of 368 graded language arts games. The mean grade at which each technique was used was determined. The means were then corrected so the lowest mean received a value of 1.0. The number values, then, do not represent specific grade levels, but rather express a relative relationship between difficulty levels.

Competition

1. Group answer, no competition: The child answers with a group of children. All children make the same response at the same time, and the group is not trying to amass score or to compete with time.
Score 1.0

2. Group answers for score: The child answers with a group, but that group is trying to achieve a score and time is not limited, i.e., the group is not trying to answer faster.
Score 2.3

3. Group answer for score and speed: The child answers with a group which is competing with another group to see who can answer best and most rapidly.
Score 2.5

4. Answers to self, no competition: The child answers alone in such a way that he is the only person aware of his answer at the time he makes it. (Teacher or another student may correct the answers later.)
Score 1.5

5. Obvious answer, no competition: The child answers alone, but in such a way that the rest of the class is aware of his answer at the time he makes it. Score is not emphasized; he simply answers a question.
Score 2.0

6. Obvious answer, self-competition: The child answers alone in front of the group. He competes with himself for a better score than he had last time or to see how far he can go in some particular task. No reference is made to scores of others.
Score 3.1

7. Obvious answer, competition for score: Child answers alone in such a way that the group is aware of his response at the time he makes it. He is competing with other members of the group for score; i.e., the child with the highest score wins. No particular reference to his past behavior except subjectively.
Score 4.2

8. Obvious answer-competition for score and speed: Child answers alone in such a way that the group is aware of his response. He competes with other members of the group for score or position and a time factor is involved in determining either score or position, i.e., he who does it fastest does it best.
Score 5.5

9. Obvious answer-competition with group responsibility score: Child answers alone in such a way that the group is aware of his answers at the time he makes them. The child is a member of a team and teams are competing for the highest score. Child cannot be helped by his teammates.
Score 6.8

10. Obvious answer, competition with group responsibility score and speed: Child answers alone in such a way that other children are aware of his answer at the time he makes it. He is a member of a team competing with another team for both score and time, i.e., fastest team with greatest number of answers wins game. Child cannot be helped by his team members.
Score 9.0

Question Props

1. Meaningful props: Meaningful props are used in asking the question, i.e., props which signify the answer: e.g., pictures or objects implying the answer in and of themselves.
Score 1.0

2. Extraneous props: Props used as a part of the question, but not meaningful to the answer, i.e., while a picture may be involved in the question, it gives no clue as to what the answer might be.
Score 1.7

3. No props: No props used in asking the question. For present purposes, flash cards are not considered props, nor are words written on the board, nor other papers having only words on them.
Score 4.1

TABLE I

CORRELATIONS BETWEEN THE GRADE LEVEL OF QUESTIONS AND THE QUESTION FORM VARIABLES
N = 172

	Competition	Question Properties	Answer Form	Answer Requirements	Total	Students' Total
Grade Level	.197**	.606**	.314**	.411**	.629**	.473**
Competition		.089	.008	-.036	.140*	.162**
Question Properties			.380**	.249**	.679**	.470**
Answer Form				.195**	.769**	.560**
Answer Requirements					.654**	.323**
Total						.634**

* $r_{.05} = .125$

** $r_{.01} = .160$

Answer Form

1. Specific object: Answer is specific. The thing called for is a picture, object, or movement. Words may be used in the question, written or oral, but child is required to find some thing or make some movement.

Score 1.0

2. Specific--part word: Answer is specified part of a word or a letter or a sound. This includes alphabetization of single words, etc.

Score 1.7

3. Specific--whole word: The answer is a specified whole word. This includes situations in which the child is required to give more than one whole word, but there is no meaningful relationship between these words.

Score 2.8

4. Specific phrase: Answer is specified phrase or number of single words which are related in meaning.

Score 4.2

5. Specific sentence: Answer is specified sentence or group of sentences.

Score 8.8

6. Non-specific: Child has freedom to make up an answer, i.e., answer not specified word, part word, or group of words. In this sense the meaning or context is the only thing specified.

Score 3.4

Answer Requirements

1. Answer present--copied--no distractors: Answer present at time child answers. Answer is to be copied or echoed. No distractors are present to divert child from answer.

Score 1.0

2. Answer present--copied--distractors: Answer present at time child answers, to be copied, echoed, or matched, but distractors are present.

Score 1.7

3. Answer present--translated--no distractors: Answer present at time child answers but must be translated from one symbol system to another with no distractors present at time child answers.

Score 2.6

4. Answer present --translated--distractors: Answer present at time child answers, and must be translated from one symbol system to another, with distractors present in terms of alternative answers.

Score 3.4

5. Answer present--reorganized--no distractors: Answer present at time child answers, and involves translation from one symbol system to another as well as organization or expression of relationships in answer itself. No distractors present.

Score 4.6

6. Answer present--reorganized--distractors: Answer present at time child answers, and involves translation from one symbol system to another as well as organization of or expression of relationships in answer. Distractors are present.

Score 6.0

7. No answer recall: No answer present at time child answers question asking for recall of a simple fact. Many completion exercises are of this type.

Score 5.8

8. No answer--reorganization: Answer not present at time child answers, and requires that he recall facts and reconstruct them or that he reorganize them in order to get the answer.

Score 7.1

The Method

The scores ascribed to each question for each of the four variables (competition, properties, answer form, and answer requirements), as well as a total sum of all the scores by the expert judge were correlated with the grade level at which the test question occurred and also with the total score determined by the student judges. The Pearson-Product-Moment Correlation Method, large sample technique, was used to determine the relationship between the variables, and further, a multiple regression equation was determined.

Results

The correlations presented in Table I indicate several factors which the designers of workbook materials use to alter the difficulty of questions. The correlation of $r = .61$ between the use of properties (in this case usually pictures) indicates a rather systematic application of pictures as a part of the question in the early grades. Recognition of items plays a more important part in the early grades ($r = .411$) and the size of the required answer tends to be somewhat smaller ($r = .314$). In spite of the obvious limitations of workbook materials on competition, the level of competitio appears to increase somewhat as the grade level of the question increases ($r = .20$).

The relation of the student judges to the variables in each case was somewhat lower than that of the expert. The large number of student judges or difficulty of the scoring system may account for the loss in accuracy.

The multiple regression equation following will allow prediction of the grade level of difficulty of the questions in terms of present practices.

$$GL = .698P + .297F + .067R + .778C - 1.26$$

Where:

GL is the grade level to which the question is assigned,

P is the index number assigned to the paragraph describing the use of properties in the question,

F is the index number assigned to the paragraph describing the form of the answer,

R is the index number assigned to the paragraph describing the requirements for the answer,

C is the index number assigned to the paragraph describing the level of competition employed in the question.

The use of this formula will produce a multiple correlation of $R_{1.2345} = .685$. This correlation affords a reasonable prediction of question difficulty in terms of current practices.

Discussion

It must be cautioned that the criteria used in this prediction formula are largely judgmental in nature, and that, while the questions themselves are varied to relate to the difficulty, this does not account for the total difficulty of the question, and that, rather, the empirical determinations which designers of workbook materials use are also in some measure based upon a change in the difficulty of the material itself. The present formula could no doubt be improved by empirical studies which hold the difficulty of the material treated in the question constant.

Summary

The relationship between the grade level assigned workbook questions in the language arts and the reading area and the variables of competition, question properties, answer form, and answer requirements were investigated. Each of the above variables significantly related to the grade level to which the question was assigned. A formula was determined for determining the difficulty level of a question, independent of the material covered in the question. Multiple correlations of all of the variables with grade level was $R_{1.2345} = .685$. It was suggested that the formula has applicability in varying the question difficulty in teaching units and for children of different I Q levels.

REFERENCES

1. Cofer, C. N. "A Comparison of Logical and Verbatim Learning of Prose Passages of Different Lengths," *American Journal of Psychology*, LIIII (1941), pp. 1-20.
2. Ebbinghaus, H. *Memory, A Contribution to Experimental Psychology*, translated by H. A. Ruger and C. Businius (New York: Teachers College, Columbia University, Education Reprints No. 3, 1913).
3. Hurlock, Elizabeth. "An Evaluation of Certain Incentives Used in School Work," *Journal of Educational Psychology*, XVI (1925), pp. 145-159.
4. Luh, C. W. "The Conditions of Retention," Psychological Monographs No. 31 (1922).
5. Lyon, D. O. "The Relation of Length of Material to the Time Taken for Learning and the Optimum Distribution of Time," *Journal of Educational Psychology*, V (1914), pp. 1-9, 85-91, 155-163.
6. Pfafflin, S. M. "Stimulus Meaning in Stimulus Pre-Differentiation," *Journal of Experimental Psychology*, LIX (1960), pp. 269-274.
7. Sims, V. "The Relative Influence of Two Types of Motivation on Improvement," *Journal of Educational Psychology*, XIX (1928), pp. 480-484.
8. Sisson, E. D. "Retroactive Inhibition; Serial vs. Random Order Presentation of Material," *Journal of Experimental Psychology*, XXIII (1939), pp. 288-294.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 5, January 1963)

The Academic Achievement of Academically Talented Students

FREDERICK R. SMITH
Indiana University

THERE HAS BEEN a notable increase of interest in recent years in special educational programs for youth of higher intellectual ability. Although those interested in these programs lament the lack of progress in this area, there is a definite trend toward making special provisions for these students in the secondary school (7).

When planning these special programs or in conducting research concerning these students, educators tend to include a larger group of students than was formerly the case. Such groups include not only "gifted" children, but also those of superior abilities who would not be included if the traditional definition of "giftedness" were employed. Conant has used the term "academically talented" to describe this greater number of students (4).

Although numerous studies have found that superior students benefit as a result of special programs (2, 3, 6, 9, 10, 11, 12, 15, 16), there are studies which do not support this point of view (1, 13, 14, 17). At best, present research concerning special academic provisions for talented students is not conclusive. Further data must be obtained before a valid appraisal of the overall effectiveness of such programs may be made (5, 8). Nevertheless, special programs for talented students in individual secondary schools are frequently based on the assumption that these programs will provide benefits not found in the regular program — especially in the area of academic achievement.

The present study was designed to consider the special programs offered by high schools in four communities. The study attempted to test the hypothesis that academically talented students enrolled in programs especially designed for academically talented students in four selected secondary schools will attain significantly higher scores on appropriate subject tests of the Sequential Tests of Educational Progress (STEP) than will academically talented students in the same schools who were enrolled in the usual college-preparatory programs.

Procedure

An academically talented student was defined as a student who is in the upper 25 percent, as measured by a general intelligence test or by a combination of the verbal and numerical scores from an aptitude battery. This is an IQ of 110 or better (108 on the Otis), or a percentile rank of 75 or above.

A questionnaire was sent to each of the 282 Michigan high schools accredited in 1959 by the North Central Association of Colleges and Secondary Schools in order to obtain information concerning special programs for academically talented students, thereby facilitating the selection of schools to be included in the study. Of the 244 schools responding, 175 reported making some special provisions for the students described.

Because of the difficulty in controlling variables which had to be considered in making comparisons of the academic achievement of pupils attending different high schools, all comparisons were to be made on an intra-school basis. It was, therefore, necessary to select schools for the evaluative phase of the study which had a sizeable number of academically talented students who were not enrolled in the special programs in their school. Thus, the experimental and control group used for each comparison would be chosen from the same school.

Four schools having special programs were selected for the evaluative phase of the study. The programs had been in operation in each of the schools for at least one year prior to the school year 1959-1960. The special programs in each of the schools were initiated in grades nine or ten, and had been in operation for only two years. Therefore, talented college preparatory junior or senior students had never been enrolled in the programs and were available for controls. These students were selected for the control groups at the time of the study by a procedure which reproduced as nearly as possible the selection of the experimental group in each subject.

A total of ten experimental groups and ten corresponding control groups were established as a result of these procedures. There were experimental and control groups from each of the four schools in English; in mathematics, comparisons were made in three schools. Comparison groups in history were established in two schools; one school had comparison groups established in science. Three types of special programs were considered in the four schools — grouping with acceleration, grouping with enrichment, and homogeneous grouping with no special methodology.

Comparisons of the Groups Prior to the Special Programs

Data common to both the experimental and control groups were analyzed to determine how the means of the groups compared in intelligence, achievement, and/or aptitude prior to the special program. These data were obtained from the cumulative record of each student. In each program, those data which seemed most appropriate for comparing the groups in the particular subject under consideration were utilized.

Comparisons of the Groups Following the Special Programs

A comparison of the academic achievement of the experimental and control groups was made following the participation of the experimental group in the special program. The Sequential Tests of Educational Progress was the achievement battery employed to measure this achievement. In each case, only the STEP tests or test appropriate for the subject included in the special program were employed. When the experimental and control groups were selected from adjacent grades, the norms for their respective grade level were utilized in comparing the achievement of the groups. The statistical procedure used in making these comparisons was the "t" test.

Findings and Conclusions

Criteria utilized by the four schools in the initial identification of talented youth included either objective criteria, subjective criteria, or a combination of the two. Each procedure was employed successfully in selecting a control group of academically talented students which compared favorably with the experimental group on abilities and achievement related to the subject area for which the identification process was intended. However, groups identified as talented in one subject area tended to be significantly different when compared on data related to achievement and/or aptitude in other subject areas. Selecting students of comparable academic talent in one subject area, and subsequently grouping them in several subjects did not appear to be effective in creating groups of comparable academic talent in all subjects.

Individual identification procedures which employed both subjective and objective criteria in selecting students academically talented in individual subject areas appeared to be the most effective and desirable identification procedures encountered in these schools.

Statistical comparisons of the achievement of the experimental and control groups, as measured by the STEP tests, led to the acceptance or rejection of the subhypotheses pertaining to the separate subject programs in each school as follows: 1) in School A, English was rejected (Table I); 2) in School B, mathematics was accepted, history was accepted with reservations, and English was rejected (Table II); 3) in School C, science, mathematics, social studies, and English were rejected (Table III); and 4) in School D, English and mathematics were rejected (Table IV). Consequently, it was necessary to reject the major hypothesis.

There was no indication that the academic achievement of academically talented students, enrolled in enriched courses designed for academically talented students in these schools, was significantly greater than that of similar students enrolled in regular college-preparatory courses in the same subjects.

Grouping of students without changing instructional techniques or subject content was not effective in increasing the academic achievement of these students when compared with the achievement of similar pupils in the regular ungrouped college-preparatory classes in the same subjects.

Accelerated courses in mathematics and world history were more effective in increasing the academic achievement of students than either enriched courses or courses which were merely grouped (Table II). The mean achievement score of the experimental group in mathematics was significantly higher at the .01 level of confidence than the mean achievement score of the control group on the same test. Likewise, on the STEP Social Studies Test, the mean achievement score of the experimental group was significantly greater at the .05 level than that of the control group. In no other comparison were the mean achievement scores of the experimental and control groups significantly different.

Discussion

This research was intended to yield data pertaining to the effect of special programs on the academic achievement of talented youth in selected public high schools. From the results of this study, it must be concluded that the achievement of the stu-

TABLE I

ACADEMIC ACHIEVEMENT OF EXPERIMENTAL AND CONTROL GROUPS IN SCHOOL A FOLLOWING SPECIAL PROGRAM

Subject	Item	Group	No.	Mean	t
English	STEP Reading Test	E	38	80.32	-.525
		C	36	82.36	
	STEP Writing Test	E	38	82.74	-.087
		C	36	83.00	

TABLE II

ACADEMIC ACHIEVEMENT OF EXPERIMENTAL AND CONTROL GROUPS IN SCHOOL B FOLLOWING SPECIAL PROGRAM

Subject	Item	Group	No.	Mean	t
Mathematics	STEP Mathematics Test	E	25	95.96	3.14**
		C	34	88.56	
Social Studies	STEP Social Studies Test	E	60	94.33	2.25*
		C	52	90.71	
English	STEP Reading Test	E	50	92.42	1.91
		C	52	88.69	
	STEP Writing Test	E	50	92.30	.626
		C	52	91.23	

** Significant at .01 Level of Confidence.
* Significant at .05 Level of Confidence.

TABLE III

ACADEMIC ACHIEVEMENT OF EXPERIMENTAL AND CONTROL GROUPS IN SCHOOL C, FOLLOWING SPECIAL PROGRAM

Subject	Item	Group	No.	Mean	t
Science	STEP Science Test	E C	18 23	88.72 93.13	-1.20
Mathematics	STEP Mathematics Test	E C	19 27	92.47 94.30	-.670
Social Studies	STEP Social Studies Test	E C	17 25	91.35 91.44	-.027
English	STEP Reading Test	E C	17 26	90.53 87.58	.828
	STEP Writing Test	E C	17 26	86.47 84.69	.389

TABLE IV

ACADEMIC ACHIEVEMENT OF EXPERIMENTAL AND CONTROL GROUPS IN SCHOOL D, FOLLOWING SPECIAL PROGRAM

Subject	Item	Group	No.	Mean	t
Mathematics	STEP Mathematics Test	E C	24 25	92.71 95.08	-1.49
English	STEP Reading Test	E C	52 45	88.21 86.62	.595
	STEP Writing Test	E C	52 45	92.29 88.87	1.56

dents in special programs in these schools as measured by the STEP tests was not significantly greater than that of similar students in the regular college-preparatory programs. This, however, is not seen as a refutation of special programs. Instructional procedures and techniques in the special programs studied, whether grouped and enriched, or merely grouped, did not differ greatly in these schools from the procedures and techniques followed in the regular college-preparatory classes in the same subjects.

This was not true of the accelerated courses in School B in which twice as much material was covered as in the regular college-preparatory classes in the same subjects. Students enrolled in these accelerated courses achieved as a group significantly higher scores than their control groups. These results seem to imply that acceleration in these schools was more effective than either grouping, or grouping with enrichment.

From this investigation, it is apparent that careful planning must precede any special program. Educators cannot assume that labeling a course as "enriched" without careful planning of this enrichment will result in a different type of classroom instruction. Likewise, it cannot be assumed that merely calling a program "special" will result in the courses in that program differing in any respect from the courses in the regular school program. It would seem a bit presumptuous to expect increased academic achievement under either circumstance.

Significant results, whether in the area of academic achievement or otherwise, can be expected only when special programs are carefully planned by administrators and teachers, involving a consideration of the nature of the pupils included in such programs and the nature of the educational program which will best challenge and stimulate these pupils.

REFERENCES

1. Abramson, David. "The Effectiveness of Grouping for Students of High Ability," Educational Research Bulletin, XXXVIII (October 14, 1959), pp. 169-182.
2. Baldauf, Robert J. "A Comparison of the Extent of Educational Growth of Mentally Advanced Pupils in the Cedar Rapids Experiment," Journal of Educational Research, LII (January 1959), pp. 181-182.
3. Bell, Mary Elizabeth. A Comparative Study of Mentally Gifted Children Heterogeneously and Homogeneously Grouped, unpublished Ed.D. dissertation (Indiana University, 1957).
4. Conant, James B. The American High School Today (New York: McGraw-Hill Book Company, 1959).
5. Educational Policies Commission. Education of the Gifted (Washington: National Education Association, 1950).
6. Engle, Thelburn L. "A Study of the Effects of School Acceleration upon the Personality and Social Adjustments of High School and University Students," Journal of Educational Psychology, XII (October 1938), pp. 523-539.
7. Goldberg, Miriam L. Research on the Gifted. Review of the literature conducted at the Horace Mann-Lincoln Institute of School Experimentation (New York 1958), mimeographed.
8. Henry, Nelson B. (ed.) Education for the Gifted, The Fifty-Seventh Yearbook of the National Society for the Study of Education, Part II (Chicago: University of Chicago Press, 1958).
9. Hollingworth, Leta S. "An Enrichment Curriculum for Rapid Learners at Public School 500: Speyer School," Teacher College Record, XXXIX (January 1938), pp. 296-306.
10. Justman, Joseph. "Academic Achievement of Intellectually Gifted Accelerants and Non-Accelerants in Junior High School," School Review, LXII (March 1954), pp. 1942-1950.
11. Justman, Joseph. "Academic Achievement of Intellectually Gifted Accelerants and Non-Accelerants in Senior High School," School Review, LXII (November 1952), pp. 469-473.
12. Keys, Noel. The Underage Student in High School and College (Berkeley: University of California Press, 1938).
13. Luchins, Abraham S., and Luchins, Edith S. "Children's Attitudes Toward Homogeneous Groupings," Journal of Genetic Psychology, LXXII (March 1948), pp. 3-9.
14. Martin, William B. Effects of Ability Grouping on Junior High School Achievement, unpublished Ed.D. dissertation (George Peabody College for Teachers, 1958).
15. Pressey, Sidney L. Educational Acceleration: Appraisals and Basic Problems, Bureau of Educational Research Monographs, No. 31 (Columbus: Ohio State University, 1949).
16. Schwartz, William P. The Effect of Homogeneous Classification on the Scholastic Attainment and Personality Development of Gifted Pupils in the Elementary and Junior High Schools, unpublished Ph.D. dissertation (New York University, 1943).
17. Shreve, Willard J. A Study of the Effectiveness of Two Types of High School Programs in Preparing Youth for Academic Success in College, unpublished Ph.D. dissertation (Ohio State University, 1957).

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 5, January 1963)

The Image of a Teachers College as Seen by High-School Counselors

SIDNEY J. DRUMHELLER
State University of New York College

AS THE MOVEMENT toward the consolidation and centralization of secondary-school facilities continues, more and more students are getting the opportunity to use the services of guidance specialists in charting their vocational and educational futures. Kenneth Cleeton found in 1951 that 3,329 students in ten Virginia colleges ranked guidance specialists fourth of 15 factors on a list of "persons from whom they received information and advice." 1*

A 15-item check list based on Cleeton's instrument was prepared and administered to 1,085 of the 1,401 students in a northeastern state teachers college, hereafter referred to as "ASTC." In the tabulation of responses to the question, "Who first started you thinking seriously about attending ASTC?", the secondary counselor ranked first. In the list of responses to the question, "Who was most influential in helping you to make your final decision to attend ASTC?", the secondary counselor ranked second to parents.

One might infer from the above information that, as the counselor describes a college and responds to questions about it raised by the counselee, he creates an image of the institution to which the student responds. The counselor's image comes into play when he decides to mention or recommend an appropriate college to a student and when he actually describes the college and its strengths and weaknesses.

The present report is concerned with three aspects of the counselor's ASTC image — its source, its accuracy, and the way it affects the counselor's neutrality and objectivity in college counseling. 2

Administering the Questionnaire

Counselors from 107 secondary schools within a 50-mile radius of ASTC were administered a questionnaire in a personal interview. The questionnaire sought information in three areas: 1) the counselors' responses to an open-ended check list concerning their sources of information and impressions about the college; 2) the counselors' responses to queries about the college typically asked by secondary students concerning enrollment, quality of faculty, students and curriculum, and other objective and evaluatable elements of their image; and 3) the counselors' statements regarding how they would differentiate in their treatment of two elementary-junior high-school teacher aspirants in each of five hypothetical counseling situations where the students differed with respect to one factor only, the other four factors being held constant. The factors considered were: differences in scholastic aptitude, in financial resources, in race, in ethnic group, and in social class. Two of the four questions posed, and the only two to be discussed here were, "Which student would you be more likely to encourage to enter ASTC in preference to a good liberal arts college with a teacher training program?", and "Which student would you be more likely to encourage to enter a good liberal arts college with a teacher training program in preference to ASTC?" Half the counselors interviewed on this part were asked to give reasons for their responses while the other half were asked if they would have responded the same way if "a state teachers college anywhere

---

* Footnotes will be found at the end of the article.

TABLE I

COUNSELORS' RESPONSES TO THE QUESTION, "WHICH STUDENT WOULD YOU BE MORE LIKELY TO ENCOURAGE TO ENTER ASTC IN PREFERENCE TO A GOOD LIBERAL ARTS COLLEGE WITH A TEACHER TRAINING PROGRAM?"

Variable Characteristic[a]	Student A	Student B	Would Not Discriminate
Scholastic Aptitude	Has high school average of 80% which seemed commensurate with ability.	Has high school average of 94% which seemed commensurate with ability.	
Percentage of Counselors Responding	40.6%	2.8%	56.6%
Financial Resources	Has three younger brothers.	Is an only child.	
Percentage of Counselors Responding	64.2%	0.0%	35.8%
Race	Negro	White	
Percentage of Counselors Responding	16.0%	3.8%	80.2%
Ethnic Group	Second generation American of South European origin.	Old stock American.	
Percentage of Counselors Responding	20.8%	0.9%	78.3%
Social Class	Upper-lower socio-economic class.	Upper-middle socio-economic class.	
Percentage of Counselors Responding	39.6%	0.0%	60.4%

[a] In each case, the four other characteristics were the same for both students.

TABLE II

COUNSELORS' RESPONSES TO THE QUESTION, "WHICH STUDENT WOULD YOU BE MORE LIKELY O ENCOURAGE TO ENTER A GOOD LIBERAL ARTS COLLEGE WITH A TEACHER TRAINING PROGRAM IN PREFERENCE TO ASTC?"

Variable Characteristic[a]	Student A	Student B	Would Not Discriminate
Scholastic Aptitude	Has high school average of 80% which seemed commensurate with ability.	Has high school average of 94% which seemed commensurate with ability.	
Percentage of Counselors Responding	0.9%	69.8%	29.3%
Financial Resources	Has three younger brothers.	Is an only child.	
Percentage of Counselors Responding	0.0%	42.5%	57.5%
Race	Negro	White	
Percentage of Counselors Responding	1.9%	14.2%	83.9%
Ethnic Group	Second generation American of South European origin.	Old stock American.	
Percentage of Counselors Responding	0.9%	26.4%	72.7%
Social Class	Upper-lower socio-economic class.	Upper-middle socio-economic class.	
Percentage of Counselors Responding	0.0%	46.2%	53.8%

[a] In each case, the four other characteristics were the same for both students.

in the state were substituted for ASTC''.

Findings

The area counselors indicated that their attitudes toward ASTC were typical of their attitudes toward the other teachers colleges of the state training elementary and early secondary teachers. The counselors ranked their five major sources of impressions about ASTC in the following order: 1)inferences made from their general attitude toward teachers colleges, 2)generalizations made about admissions standards derived through routine contacts with the admissions office, 3)use of the literature distributed by the college, 4)personal contacts with the college through an official college representative, and 5)reports from professional acquaintances who were ASTC graduates.

The counselors were most accurate in assaying the scholastic aptitude of the ASTC freshmen and the academic achievements of the college faculty. When they erred, it was in the direction of characteristics which have not existed for several years – i. e., compared with the present situation they tended to perceive the student body as consisting of fewer students, a greater percentage of women, a greater percentage of commuters, and a greater percentage of self-supporting students; and they tended to see the curriculum as more professional, less liberal, and more of a vocational dead end.

Seventy and eight-tenths percent of the counselors received their undergraduate degrees from private colleges and 75. 5 percent did all their graduate work in private colleges. The images held by these graduates were remarkably similar, however, to those of the counselors who had attended teachers colleges.

Sixty-seven and eight-tenths percent of the responses in the hypothetical counseling situations indicated that the respondents would not differentiate in their treatment of the students presented. However, the remainder of the responses indicated that many of the counselors have a strong tendency to perceive ASTC as more appropriate for those prospective academic teachers with low-status characteristics and the good liberal arts college with a teacher training program as more appropriate for those students with high-status characteristics. That is, they tend to encourage toward the teachers college the student with the 80 percent high school average rather than the student with the 94 percent average, the student with fewer financial resources, the Negro student rather than the white student, the student with recent South European origins rather than the old stock American, and the student of the upper-lower socio-economic class rather than the student of the upper-middle socio-economic class (see Table I).

The opposite tendency exists in encouraging students toward the good liberal arts college (see Table II). Although the data do not show the degree to which the counselors actually influence students in these directions, they point out that the counselors do discriminate and that they are aware they do. Actually, 91. 5 percent of the counselors indicated that they would differentiate in their treatment of the students presented in at least one area.

Summary

One hundred and seven counselors in secondary schools within a 50-mile radius of ASTC were interviewed regarding their image of the teachers college. More than 91 percent indicated they would, under certain circumstances, differentiate in their treatment of aspiring teachers regarding whether they would encourage them toward a liberal arts college or ASTC. In general, they were more than 20 times more likely to encourage the student with the high-status characteristic to attend the liberal arts college than ASTC and more than 20 times more likely to encourage the student with the low-status characteristic to attend ASTC than the liberal arts college.

The counselors' image of ASTC seemed to be based on a stereotype of a teachers college, tended to be typical of the counselors' image of the other teachers colleges in the state, and where its aspects could be evaluated, tended to be from five to ten years outdated.

FOOTNOTES

1. Kenneth Henry Cleeton, An Analysis of the Factors Which Influence Students When They Choose a College, unpublished doctoral dissertation, University of Colorado (Boulder: 1951).
2. For a complete report of the study, see Sidney John Drumheller, The Image of a Teachers College as Seen by High School Counselors, unpublished doctoral dissertation, Teachers College, Columbia University (New York: 1961).

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 5, January 1963)

Evaluation Based Upon Observation of Reports of Supervisors of Student Teaching

RACHEL D. WILKINSON
Mount Vernon, New York

A STUDY of observation reports made by supervisors of student teachers can reveal some interesting facts about an institution's program for student teaching. The purpose for making such a study may be to point up necessary changes in the present practices, as well as to emphasize practices which seem reasonable to follow.

An example of such an evaluation will be described to point out possible areas to be studied or appraised. This study covered a period of five months; i. e., one term, in which 32 students were assigned to student teaching in public and private elementary schools in the New York City vicinity. The students were, for the most part, graduates of liberal arts institutions and were working for their first graduate degree. They were observed by the members of the university staff who presented written reports of their observations and conferences with the students. These reports became a part of the students' records at the institution.

The report form, an 8 1/2 x 11 mimeographed sheet, contained spaces for identification data, and for free comments. The identification data included: name of student, data and time of observation, school grade, lesson topic. There were three spaces provided for comments: 1) supervisor's comments, 2) summary of strengths, and 3) suggestions for improvement. These data from the reports have been summarized and used as a basis for the analysis in this study.

The summary of the supervisors' reports is here presented in five sections: 1) distribution of the student teaching experience, 2) the variety of experience, 3) the areas in which the students worked while being observed, 4) the number of times observed, and 5) strengths and weaknesses of student teachers. A summary of the findings with regard to each of these items will be given followed by questions and general suggestions for future evaluations.

Distribution of Levels of Experience

The first item reviewed was the distribution of levels of experience of the 32 students. It was found that 13 of them had experiences in the lower grades, that is from the kindergarten through the third; 10 had experience in the upper grades, grades four through six; and nine had experience on both levels. In other words, 23 had experience in only the lower or upper grades, and only nine had experience on both levels.

Within the distribution described in the previous paragraph, 11 students remained with the same class for the five-month period. Four of these were in the lower grades (2 in the kindergarten, 1 in the first grade, and 1 in the third grade); and 7 were in the upper grades (2 in the fourth grade, 2 in the fifth grade, 3 in the sixth grade).

Therefore, in view of the distribution of experiences for the 32 students, the following questions might properly be raised: Is it desirable for all students to have experience on both levels? Is it more desirable for students to remain with the same class throughout the student teaching experience? Or is it preferable to have a flexible program which fits the needs of the individual students? Could it be that administrative policy in some schools prevented students from having experience on more than one level, or with more than one class?

The Variety of Experience

The variety of experience[1]* of the students was summarized as the second item of the survey. Eleven of the 32 students, as already explained, received experience in only one grade. Eleven others received experience in two grades during the five-month period, while seven had three different experiences, and three had more than three experiences. The following list will show the variety of experience for the 32 students by grade level.

Grade	Total Having Experience On This Level
K	7
1	13
2	10
3-4	1
4	8
4-5	2
5	6
6	7
7	1

---

* Footnotes will be found at end of article.

This chart shows that the first and second grades led all others as places where experience was received.

The following questions might be raised about the variety of experience for the 32 students: Why were the lower grades used more often than the others? Does this indicate a trend in preparation preference? Does this indicate that students feel more comfortable in these grades? Could it be that supervisors request placement in the lower grades above all others, or is this a demand of the school administrators?

Curriculum Areas Observed

The supervisors made 116 visits to observe the 32 students during the term. To answer the question, what did the student teachers teach for these visits, a compilation of the curriculum areas in which the students were teaching when observed shows that most of the teaching was done in six curriculum areas. These were language arts (34)[2], mathematics (19), reading (18), science(12) social studies (9), and art (7). There were a few lessons presented in music (4) and health (3). One person had charge of morning routines, and another was giving a test when the supervisor came to observe.

The matter of which curriculum area received more experience raises these questions: Should more emphasis be placed on music and health? Why are these areas more or less neglected? Is it important to observe a student in roles other than that of teaching subject matter? Should the bulk of the experience be required in language arts, mathematics and science? Who should make the decision as to the area in which the student is to be observed? Should a student be observed in a different area each time? If a student is weak in a certain area, should he be required to do more teaching in this field?

A list of some of the topics of the lessons were observed is given below:

Language Arts

Handwriting
Use of comma
Use of capitals
Word meanings
Synonyms and antonyms
Pronouns
Use of apostrophe
Spelling of Desert words
How animals get ready for winter
Creative writing
Writing a letter
Good conversation
Introductions
Safety education
Discussion of story

Reading

Two groups
Experience charts
Silent and oral reading for slow readers
Silent and oral reading for superior readers
Comprehension
New vocabulary
Phonics
Fluency and interpretation
Readiness-difference in pictures
Oral reading
Story telling
Story of Sleeping Beauty
Story-Circus in the Jungle
Descriptive words
Discussion of the circus
How to select and evaluate books
A walk in the park
Show and tell
Initial consonants
Phonics
Spelling test

Music

The Noble Duke of York
Thanksgiving
Concept of Octave

Health

What foods are good for us
Games

Social Studies

Growth of big business
Trip discussion
Sahara Desert
Columbus
The Lapps and Eskimos
Completing a directional map
Review of Middle Atlantic states
Sugar
Henry Hudson
Mexico

Art

Arts and crafts
Cut paper pictures
Making Pinocchio puppet
Construction of hand puppets
Making a paper fruit bowl
Paper and its properties
Making a sock doll

Mathematics

Money
Making change in U.S. currency
One less than
Telling time
The clock
Review
The calendar
Test
Number concepts
Concept of 7
Concept of 3
Concept of 5
Fractions
Fractions of a whole number
Three-fourths
Adding by endings
Grouping
Workbook drill

Science

Magnets
What magnets can do
The compass and magnetism
Magnetism and the compass
Electromagnets
Air takes up space
Friction
Our teeth
Birds
Milkweed plant
Foods

Miscellaneous

General testing
Morning routines

Individual Observations

A look at the number of times each student was observed by his supervisor shows that most of the students were observed three or four times each during the term. However, five were observed five times each, and one person was observed six times. The chart on the next page summarizes the number of observations per student.

No. of Observations Per Student	No. of Students
1	1
2	2
3	11
4	12
5	5
6	1
Total	32

A summary of the number of different curriculum areas in which a student was observed shows that all but 10 of the 32 students were observed teaching in three or more different curriculum areas. Nine students were observed in two different areas, 12 in three different areas, eight students in four different areas, and five in two different areas as follows:

No. of Different Curriculum Areas in Which Observed	No. of Students
1	1
2	9
3	12
4	8
5	2
	32

In view of the survey of curriculum areas in which student teachers had experience, the following questions might be asked: Is it important for students to be observed in more than three curriculum areas? How often should the supervisor observe the student? What factors decide the policy about the number of observations per student? To what extent should grade-level placement determine the curriculum areas to be taught by the student?

Strengths of Student Teachers

In their reports the supervisors listed the strengths and weaknesses which they observed while each of the 32 students was teaching.

The strengths seem to involve personal attributes, manner and attitudes of the students, evidence of planning, and the teaching procedures. A glance at the items in these various categories would seem to indicate what the supervisors considered desirable traits for teaching.

With regard to personal strengths, pleasing personality, appearance, and voice were the main phases listed. The personality was described as pleasing, pleasant, fine person, or simply the word "personality" was used. Appearance was described in the following terms: good appearance, attractive appearance, neatness, dress, attractive and friendly, and simply the word "appearance." The word "voice" was used alone to describe strengths along with "firm but friendly voice, voice and speech, tone of voice, good voice, and voice well modulated.

Some of the strengths depicting desirable manner and attitude were mentioned more often than others: mature, sincerity, enthusiasm, and rapport. Specific wording of these attributes is listed below, with similar items grouped together:

1. Cheerful, sense of humor, calm and happy, ability to joke, happy about work
2. Positive attitudes, use of positiveness and courteous manner, use of praise
3. Sweet and gentle, manner is gentle and assuring, gentle manner
4. Mature, poised, patient
5. Alert, industrious, willing, competent, capable, cooperative, works hard, seriousness toward work, hard worker, conscientious, willing and good natured
6. Manner, natural manner, sincere, professional manner, punctual
7. Enthusiasm for teaching, enthusiasm, curiosity about teaching
8. Gets along well with superiors
9. Likes children, respect for children, comfortable with children, and they with her, manner with children, well liked by children, understanding and sympathy, superior understanding, desire to work with children, rapport, ease with children
10. Patient, wonderful with children, warmth toward children
11. Willingness to accept criticism and learns quickly, willingness to learn, willing to follow suggestions, willingness to take suggestions.
12. Self-evaluation, awareness of difficulties

With regard to the background of the students, the strengths noted included these: superior understanding of children, good background, writes well, artistic ability, expert pianist, command of subject matter, good at record keeping.

The descriptions involving planning for teaching indicated that the supervisor liked the planning techniques, careful planning, thorough planning, selection of teaching material, and charts and attractively prepared teaching aids.

The supervisors described the observed strengths in teaching as follows:

Excellent teaching, splendid, superb lesson, outstanding lesson

Fine motivation, creative, resourceful, original and enthusiastic

Use of illustrative material, use of objective material

Routines, moved about, transition from one activity to another, thorough, reviews

Timing, alert to need for change of pace, adjustment to situation

Questioning, explanations

Ability to handle a class, no trouble with discipline

Character development, use of praise, mature judgment, sensitive to emotional problems, recognition of needs of children

Participation of children

Board writing clear, no errors

Weaknesses of Student Teachers

The improvements which the supervisors suggested that the student teachers needed to make were almost the opposite of the strengths which were observed by them. The improvements needed may be classified as weaknesses in personal qualifications, background, manner, and teaching techniques.

The comments with regard to weaknesses in personal qualifications included: untidy appearance, overweight, poor health, and inadequate speech or voice defects.

The backgrounds were inadequate due to lack of information regarding teaching methods, poor handwriting, and a need of insight into how children learn.

Manner and attitude showed a lack of self-confidence, an unwillingness to accept suggestions, lack of poise and maturity, tardiness, irregular attendance, and inability to adjust to the children and to peers.

The complete list follows:

Manner, immature attitude, lacks spark and vitality

Lacks confidence, not assertive enough, negative approach, authoritarian, does not relax, tense, not sure of herself, feels inferior and lacks self-confidence, does not assume responsibility

Not punctual at school, attendance irregular, personal problems and illness interfered

Criticizes other teachers, an "apple polisher"

Unwilling to follow suggestions, unwilling to accept criticism

Sarcastic manner with children, emotionally involved with students

Too senstiive and reads trouble into situations where none exist

The weaknesses in teaching seemed to stem from these inadequacies: preparation for the lessons, use of teaching materials, motivation techniques, sense of timing, skill with questioning, routines, and adjustment to class and to individual children. The complete list of needed teaching improvements follows:

No plan, planning inadequate, duplicated material would have helped, does not plan for slow and bright students.

Tried to cover too many concepts in one lesson, poor timing, lesson too long, might use duplicated material to save time

Need to use concrete materials, techniques and methods, teaching skills, techniques of teaching

Does not know how to motivate, poor motivation, no ingenuity

Not challenging, does not require children to think

Directions not clear, questioning skills, questions and comments not suited to age level, vocabulary above children, repeats answers of children, does not allow children to talk enough, repeats what they say

Routines, untidy with classroom routines, does not adapt to rules and methods of class, attention to details, writes too low on board, careless habits with board work, overuse of O. K.

Participation of children, no participation by some children, needs to involve children in process, called on some children too often, some children too soft spoken

Did not have children put away objective material before going on

Discipline, getting attention of whole class, lets other groups get restless, does not get total attention, starting lesson without getting attention at the beginning

Insists on formal behavior

Maintains stationary position

The writer feels that the factors included under strengths and weaknesses found by the supervisors working in one term with 32 student teachers are revealing. Perhaps such a list might be made available to students as a check-list for self-evaluation. If over the years the same weaknesses are apparent in each group, it may mean that the screening process for the selection of student teachers is inadequate.

It was most interesting that the supervisors involved used similar terminology in describing the strengths and weaknesses of the student teachers. These comments were submitted without any previous discussion of what should be listed. Each comment was therefore an individual matter on the part of each supervisor.

FOOTNOTES

1 The variety of experience was determined by the placement of the student at the time the supervisor observed the teaching. The supervisors observed each student before a change of grade was made.

2 The figures in parentheses show the number of lessons observed by the supervisors in the area during the term. The reading might be combined with the language arts, but they were separated in this report since so many reading lessons were given.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 5, January 1963)

The Prediction of Success in Initial College Mathematics Courses*

LOREN V. COROTTO
Napa State Hospital
Imola, California

ALTHOUGH EVIDENCE is somewhat contradictory, aptitude tests and achievement tests designed to measure specific content areas appear to provide greater predictive accuracy of success in mathematics than do tests of general intelligence or ability. Seigle (7) reports a study of the prediction of success in college mathematics at Washburn University. Results indicate that at Washburn, an entrance mathematics test was the best single predictor for success in algebra courses. When Krathwohl (5) compared success in college algebra with scores on the Iowa Mathematical Aptitude Test, he found a significant correlation. Wallace (9) found the Iowa Foreign Language Aptitude Test and the Mathematics Placement Test (form B-2) to be significantly more effective in predicting grades for mathematic courses than the ACE for three levels of algebraic difficulty. The purpose of this paper is to present large sample data on the relative predictive effectiveness of a test of general scholastic aptitude and of a test of mathematics achievement.

Methods and Procedure

In the summer of 1957, the Registrar's Office of the University of Houston, at the request of the Counseling and Testing Service, made photostatic copies of the permanent record cards of all the students who entered the University during the fall semester of the 1955-56 academic year. A thorough examination of these records was made to select all first-semester freshmen whose initial mathematics course in college was in one of a series of three algebraic courses having different levels of complexity. Although the majority of the students took their initial mathematics courses during their first semester in college, any student who completed one of these three algebraic courses as his initial mathematics course during the period from September 1955 to June 1957 was included in the study. Each student had to have received a final grade of A, B, C, D, or F in his initial mathematics course in order to be included in the study.

Subjects

A total of 513 students — 41 females and 472 males — met the requirements for inclusion in the present study. Ages of the subjects ranged from 17 years to 40 years, with a median age of approximately 18 years.

Description of Mathematics Courses

A description of the mathematics courses explored in this study is presented in the University of Houston's catalog for the 1955-56 academic year.

Mathematics 030: Beginning Algebra. A beginning course in algebra designed for those students who have had no former work in algebra. This course carried terminal credit only. Junior College aid will not apply (University of Houston, p. 86).

Mathematics 153: College Algebra. A review of high-school algebra followed by the standard work in college algebra. Five lecture hours per week. Three hours degree credit and two hours terminal credit. Prerequisite: One unit high-school algebra or Mathematics 030 (University of Houston, p. 86).

Mathematics 131: College Algebra. Quadratic equations, simultaneous quadratic equations and inequalities, progressions, the binomial theorem, determinants, elements of theory and equations, and partial fractions. Formerly offered as Mathematics 133. Prerequisites: Two years of high school algebra and approval of department chairman (University of Houston, p. 86).

---

* This article is based on data from an unpublished dissertation submitted to the Graduate School of the University of Houston in partial fulfillment of the requirement for the degree of Doctor of Philosophy.

Description of Instruments Used

The two test instruments utilized in this study are the American Council on Education's Psychological Examination for College Freshmen (ACE), and the Mathematics Screening Test (MST), constructed and developed at the University of Houston.

American Council on Education Psychological Examination for College Freshmen. This scale of mental ability and general scholastic aptitude yields three scores: the linguistic score (L), the quantitative score (Q), and the combined L and Q scores produce the total score (T).

Mathematics Screening Test. The MST was jointly developed by the Mathematics Department and the Counseling and Testing Service of the University of Houston. The test is composed of 35 algebraic problems arranged in order of increasing difficulty. The test is timed, and in such a manner that speed is a factor in a student's score. The final score is the total number of problems in the test that have been answered correctly. The distribution of the 513 MST raw scores obtained by the subjects rather closely approximates a normal curve (Mean = 15.01, SD = 8.26), although there is a slight positive skewness.

Criterion Measure. The criterion measure used in the present study is the mathematics grade achieved by each student in the sample. The grade distributions of Mathematics courses 030, 153, and 131 approximate a normal curve, and chi-square comparisons of the distributions indicate that the distributions were comparable; i.e., they were not significantly different. Initial success was operationally defined as a grade of A, B, or C, while failure was operationally defined as a grade of D or F.

Analysis of the Data

All correlations in the present study are computed from the raw scores obtained on the ACE and MST. The results are presented in Table I. Significant correlations were revealed between the total scale of the ACE and the pass-fail criterion of all Mathematics 030 students as well as all students in Mathematics 030, 153, and 131 combined. The qualitative scale of the ACE correlated significantly with success and failure in Mathematics 030 as well as with the combined Mathematics 030, 153, and 131 distribution.

There were significant correlations between the MST and success in specific mathematics courses. The results are shown in Table II. Biserial correlations of the MST scores with success in Mathematics 153 and Mathematics 131 are significant at the .01 level of confidence. There is not a significant correlation coefficient between MST scores and successful completion of Mathematics 030, although the ACE did yield significant correlations with Mathematics 030.

Discussion

It is readily apparent from a glance at Table I that the ACE scores do not distinguish between successful and unsuccessful students whose initial mathematics course is Mathematics 153, or Mathematics 131. The significant correlations between the quantitative scale and the total scale of the ACE, and success in Mathematics 030 and the total group, do not indicate a degree of relationship that would warrant its use for individual prediction, although it might possibly contribute towards gross-group prediction and differentiation. There are several possible explanations for the significant correlations between the quantitative scale and the total scale of the ACE and success in Mathematics 030. Since Mathematics 030 is purely a remedial course, there is a good probability that the content of Mathematics 030 is more similar to the content of the ACE than is the content of either Mathematics 153 or Mathematics 131. It seems quite likely that variance associated with variables such as content, teaching methods, and mathematical background might tend to minimize the importance of the linguistic ability as measured by the linguistic scale of the ACE.

The failure of the linguistic subscale scores to correlate significantly with success in the mathematics courses examined in this study is at variance with the results of studies by Berdie, Dressel, and Kelso (1), and Hoerres and O'Dea (4).

The MST correlates significantly with Mathematics 131 and Mathematics 153, while failing to correlate significantly with success in Mathematics 030. The results indicate that the MST is a much better instrument for predicting success in Mathematics 153 and Mathematics 131 as an initial mathematics course than is the ACE. Success in either Mathematics 153 or Mathematics 131 appears related to mathematics achievement much more strongly than it is to general intelligence or general academic aptitude, at least as measured by the ACE. These results are similar to the findings of Bruce (3), Seigle (7), Bromley and Carter (2), and Remmers, Elliott, and Gage (6).

TABLE I

BISERIAL CORRELATION VALUES BETWEEN ACE SCORES AND PASS-FAIL STATUS IN INITIAL MATHEMATICS COURSES

	Mathematics Course											
	030 (N = 127)			153 (N = 202)			131 (N = 142)			All Courses (N = 471)		
ACE Scale	$r_{bis.}$	SE	Number Passing	$r_{bis.}$	SE	Number Passing	$r_{bis.}$	SE	Number Passing	$r_{bis.}$	SE	Number Passing
Total	.28*	.10	76	.08	.09	124	.16	.11	96	.17*	.06	296
Linguistic	.17	.10	76	.07	.09	124	.11	.11	96	.11	.06	296
Quantitative	.31*	.10	76	.14	.09	124	.14	.11	96	.20*	.06	296

* *Significant at the .01 level for a two-tailed test.* The unstarred correlational values failed to reach the .05 level *of significance for a two-tailed test.*

TABLE II

BISERIAL CORRELATION VALUES BETWEEN MST SCORES AND PASS-FAIL STATUS IN INITIAL MATHEMATICS COURSES

Mathematics Course	*Total Sample*	$r_{bis.}$	$SE_{r_{bis.}}$
All 030 students	135	.14	.11
All 153 students	220	.42*	.07
All 131 students	158	.40*	.09

* Significant at the .01 level for a two-tailed test.

Summary

The present study was concerned with exploring the relationships between 1) success in an initial mathematics course and scores on the three scales of the ACE, and 2) success in an initial mathematics screening test developed at the University of Houston. The results show that the ACE is a better predictor for Mathematics 030, a purely remedial course, while the MST is a better predictor of initial success in Mathematics 131 and 153 (courses of college level algebra).

REFERENCES

1. Berdie, R., Dressel, P., and Kelso, P. "Relative Validity of the Q and L Scores on the ACE Psychological Examination," Educational and Psychological Measurement, XI (1951), pp. 803-812.
2. Bromley, A., and Carter, G.C. "Predictability of Success in Mathematics," Journal of Educational Research, XLIV (1950), pp. 148-150.
3. Bruce, W.J. "The Contribution of Eleven Variables of the Prognosis of Academic Success in Eight Areas of the University of Washington," Dissertation Abstracts (Seattle: University of Washington, 1953).
4. Hoerres, M.A., and O'Dea, J.D. "Predictive Value of the ACE," Journal of Higher Education, XXV(1954), p. 97.
5. Krathwohl, W.C. "Prediction of Average Class Achievement by Means of Aptitude Tests," Journal of Engineering Education, XXXVII (1946), pp. 234-242.
6. Remmers, H.H., Elliott, D., and Gage, N. "Curricular Differences in Predicting Scholastic Achievement: Applications to Counseling," Journal of Educational Psychology, XL (1949), pp. 385-394.
7. Seigle, W.F. "Prediction of Success in College Mathematics at Washburn University," Journal of Educational Research, XLVII (1954), pp. 577-588.
8. University of Houston: The College of Arts and Sciences, Undergraduate and Graduate Catalog, 20-B (Houston: University of Houston).
9. Wallace, W.L. "The Prediction of Grades in Specific College Courses," Journal of Educational Research, XLIV (pp. 587-597).

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 5, January 1963)

Evaluation of a College Study Habits Course Using Scores on a Q-Sort Test as the Criterion

JAMES W. CREASER
University of Illinois

ENTWISLE (2), in a review of the literature, found 22 studies dealing with the evaluation of study habits courses. Twenty of these used grades in school or college as a criterion by which to measure the value of the course. About half of the studies used some measure of reading proficiency as a second criterion.

Only one of the articles reviewed by Entwisle, that of Tresselt and Richlin (4) used study habits as a criterion. They obtained rather large increases between a pre- and post-test with the Wrenn Study Habits Inventory.

The present study investigates changes in study habits and attitudes during a study habits course, using a Q-sort test as a criterion. This test includes items which describe study skills, attitudes about teachers and college, and satisfaction with progress and goals.

Procedure

The items for the test were put on cards, and in taking the test, the student read each item and then put the card into the appropriate pile. There were seven piles, and the student put the items which best described him in pile number one, and the items which least described him in pile number seven, with most of the items on the continuum between the two extremes.

The distribution of cards among the piles was prearranged so that the items formed a normal distribution, and constituted the student's profile. This profile was then scored by correlating it with the "ideal" profile which was constructed by the counselors.

The students took the test on the first day of the study habits course, with half of them taking Form A, and half Form B.

The post-test was administered on the next-to-the last day of the course, with each student being tested on the form opposite from the one he had originally taken. Thus, each student had two scores, one obtained at the beginning, and one at the end of the course.

Results

The correlations were transformed to Fisher's z-scores, according to McNemar (3), so that they might be averaged. The results in Table I are in z-scores.

The results show that a highly significant change in the mean value of the scores occurred during the course. The average student profile on the post-test was more nearly like the "ideal" than on the pre-test. This indicates that the students, as a group, showed a sizeable improvement in their study habits and in their adjustment to college.

Control

It is certainly true that many students improve their study habits and adjust to college without help from group or individual counseling. Thus, the possibility that this group of students did not solve any more of their problems with the group help than they would have without it should be controlled.

For this purpose, the students were divided into two groups and classified as "old" or "new" students. The "old" students were those with one to four (average of 1.4) previous semesters of college work. The "new" students were in their first semester at college. The pre-test scores of these two groups were compared to measure the effect of previous college experience upon the kinds of problem areas covered by the test. The comparison is shown in Table II. 1*

The results of this comparison show that the

---

* Footnotes will be found at the end of the article.

TABLE I

COMPARISON OF PRE- AND POST-TEST SCORES FOR STUDENTS TAKING THE STUDY HABITS COURSE

(The N for this comparison is 88. The correlation between the pre- and post-test was .29).

	Pre-test	Post-test	Difference Between Means
Mean	.04	.26	.22*
σ	.25	.26	
σ_M	.026	.029	

* Significant at .001 level.

TABLE II

COMPARISON OF PRE-TEST SCORES FOR "NEW" AND "OLD" STUDENTS

	New Students	Old Students	Difference Between Means
N	197	56	
Mean	.04	.04	.00*
σ	.25	.24	
σ_M	.018	.032	

* Not significant.

students with previous college experience had, as a group, the same average score on the pre-test as did the group of beginning students. This would suggest, although not prove, that the previous exposure to college had not had an effect upon the improvement of their study habits and attitudes.[2]

During the course, the increment in the average score for the "old" students was .22 z-scores, the same as for the group as a whole.

Discussion

The results of Entwisle's review may reflect a feeling among the investigators that improved grade average is the most important benefit which a study course offers. However, it seems to this author and to some of his colleagues that this is a rather narrow criterion by which, alone, to measure the success or failure of the study habits program.

The study habits course at the University of Illinois, Chicago Undergraduate Division, is usually conducted as a semi-structured group-counseling program. This allows for considerable discussion by the students of the varied problems with which they may be faced. And it has often been noted by the counselors that the students come in with, and often solve, problems which are not necessarily directly related to grades.

For example, one student decided, after discussing it in the group, that he was going to change his curriculum from Pre-medicine to Business Administration. His grades did not improve as a result of this decision, in fact, they decreased slightly. But he did get rid of his ulcer.

Further, in a factor analytic study of the Q-sort test, the author (1) described five factors which typified various student types. It seemed from this that certain students had problems about "school work" but not with "grades". One of the factor types, the "diligent student", apparently made acceptable grades, but was interested in learning study methods which would help him cut down on the prodigious amount of time he was spending in study.

Then we have students' comments. One said, "I enrolled because I have always had trouble talking in class. I think this group will help me overcome that."

These and other observations indicate that while improved grades may likely be the goal of some of the students, it is not necessarily the goal of all. It would appear that it is important to use multiple criteria when evaluating a study habits course. And it would also seem that one of the criteria should be some measure of the student's satisfaction with the adjustment he has made to studying and college life.

The Q-sort test used in this study gets at many things which should result in improved grades, and it also covers many areas which don't relate directly to grades, but do relate to personal adjustment to college. It is felt that both should be part of a well-rounded study habits program.

Summary

The results of using a Q-sort test as a criterion for evaluating a study habits course are reported. The general evaluation of such courses is discussed, and it is proposed that grades alone are not an adequate criterion.

FOOTNOTES

1. The N for this comparison is 253 instead of 88, as it was for the pre- and post-test comparison. There were 253 students who took the pre-test. However, some of them withdrew from the course or were absent the day the post-test was given, and some of the counselors did not give a post-test to their groups. Thus, out of the original N of 253, there were 88 who had both pre- and post-tests.
2. These conclusions, of course, are based on the students who took the study habits course. It should not be inferred from this that the same would be true for students in general.

REFERENCES

1. Creaser, J. "Factor Analysis of a Study Habits Q-Sort Test," Journal of Counseling Psychology, VII (1960), pp. 298-300.
2. Entwisle, D. "Evaluation of Study Skills Courses: A Review," Journal of Educational Research, LIII (1960), pp. 243-251.
3. McNemar, Q. Psychological Statistics, John Wiley and Sons, Inc. (New York, 1949).
4. Tresselt, M., and Richlin, M. "Differential Prognosis in a College Study Methods Course," Journal of Psychology, XXXI (1951), pp. 81-89.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 5, January 1963)

A Method for Identifying Prospective Counselees in College

ROBERT M. ROTH
Hampton Institute

THE PURPOSE of this study was to test the practicability of a particular method for identifying the students who are prospective counselees. The areas of adjustment and anxiety were selected for measurement because it was assumed that adjustment and anxiety are related to students seeking counseling. It is assumed further that a combination of poor adjustment and high anxiety is a condition related to the seeking of counseling. A test of adjustment, the California Test of Personality (2), and a test of anxiety, the Manifest Anxiety Scale (1) were selected because of ease of administration to large groups. It would seem reasonable that the combination of poor adjustment and high anxiety would be reflected in the test results of students who did seek counseling at a counseling center.

The hypothesis which this study attempts to test is as follows:

> There will be a significant difference between the expected scores on the California Test of Personality and the Manifest Anxiety Scale and the scores made by those students who seek counseling. These scores will tend towards the combination of greater anxiety and poorer adjustment for the group seeking counseling.

Procedure

The California Test of Personality and Taylor's Manifest Anxiety Scale were administered to the entire freshman class. The total scores on the California Test of Personality and the Manifest Anxiety Scale were culled for students who sought counseling in the Office of Psychological Services from September, 1958 to April, 1960. This group consisted of 18 males and 27 females. The scores of the entire freshman class were normalized into categories of excellent (first quarter), average (second and third quarter), and critical (fourth quarter) for both measures. The frequencies in the categories of scores of the students seeking counseling were compared with the expected frequencies based on the performance of the freshman population. A chi-square analysis using Yates' correction was applied to the date.

Results

Table I contains the observed and expected frequencies in scoring categories for those freshmen seeking counseling. Using Yates' correction, a chi-square value of 17.0 was determined. This value was significant at .01 level of confidence for four degrees of freedom. These results offer evidence that a combination of poor adjustment and high anxiety as measured by these tests was found in students who sought counseling.

Conclusion

This study has indicated that poor adjustment and high anxiety as reflected in the selected measures are related to students' seeking counseling. Inherent in this study is the implication that there are other factors involved in the students' motivation for seeking counseling, and that the group of students who actively seek counseling belong to a population who are prospective counselees. The method presented in this study has been shown to be a practical technique for identifying this population.

REFERENCES

1. Taylor, Janet, "A Personality Scale of Manifest Anxiety," *Journal of Abnormal and Social Psychology*, XLVIII, pp. 285- 90.
2. Tiegs, Ernest; Clark, Willis; and Thorpe, Louis. *California Test of Personality, Secondary Form*, California Test Bureau (1953)

TABLE I

CHI-SQUARE CONTINGENCY FOR SCORES MADE BY FRESHMEN ON THE CALIFORNIA TEST OF PERSONALITY AND MANIFEST ANXIETY SCALE MEASURES

California Test of Personality	Manifest Anxiety Scale Obtained Frequencies			Expected Frequencies		
	Critical	Average	Excellent	Critical	Average	Excellent
Excellent	0	3	4	1.4	8.6	6.3
Average	7	11	1	4.5	7.9	7.7
Critical	12	6	1	6.3	2.7	0.4

$X^2 = 17.0$ (significant at the .01 level of confidence)
df = 4

JeR Field News

THE NEW AUGMENTED ROMAN ALPHABET for teaching reading is being used for the second year with 1500 English children. This alphabet has 43 letters, 24 of the original 26 and 19 new ones. This is not a revised spelling plan, but is an effort to see if children can learn to read more easily if the spelling problems are postponed.

PHI DELTA KAPPA HAS JUST PUBLISHED "Dissemination and Implementation of Educational Research". This publication, based on a two day symposium, covers the functions of educational research in the areas of electronics, philanthropic organizations, study councils, the USOE, the communications media, etc. (208 pp., $3.00, from Phi Delta Kappa, 8th and Union, Bloomington, Indiana.)

THE STATE DEPARTMENT OF EDUCATION of California recently completed a study of pressures on children and youth. The pressures found were: 1) fear of the future, 2) demands to get ahead academically, 3) unrealistic demands for scholastic achievement, 4) demands to push advanced curriculum in the lower grades, 5) demands for too much homework, 6) competitive race for college, 7) demands to get ahead socially, 8) overcrowding the lives of children, 9) conflicting values, 10) questionable school practices.

OVER 40,000 AMERICAN SCHOOLS have no library, although their enrollment is over 10 million pupils.

EDUCATION IS ONE OF THE SIX MAIN AREAS which the Ford Foundation will strongly support in the 1960's. In educational research the Fund plans to help:

1) Research and experimentation in learning and teaching techniques, such as team-teaching, flexible scheduling of classes, and programed learning. The Fund will encourage the use of non-commercial television for educational purposes and will help education to incorporate the continuing stream of educational innovations into comprehensive improvement programs covering the entire school systems.

2) To develop foreign areas in international studies as strong integral features of university teaching and research. To strengthen the ability of undergraduate faculties, to incorporate non-Western and other international studies into liberal education, and to improve the educational and living experiences of foreign students in the United States. To improve the teaching of foreign languages, including the science of linguistics, and to develop American training and research resources for the teaching of English abroad as a second language.

THE OHIO BUREAU OF UNEMPLOYMENT COMPENSATION reports the relationship between education and employment, and the corresponding public cost of inadequate education. The Bureau found that of 215,000 unemployed men and women in Ohio, two-thirds of them had not completed high school, while just over one percent were college graduates. In the best educated age group, 20-24 years, four out of ten had failed to graduate from high school.

A STUDY OF SCHOOL BOARDS was recently made by Alpheus L. White, USOE specialist on school boards. From a survey of more than 4,000 school districts having 1200 or more pupils, he found that most school board members are elected by popular vote; most have five or more members; 47 percent had been in office five years or more; almost half were college graduates; almost all boards held open meetings, but few citizens attended; about 60 percent of the boards had written policy manuals; only ten percent of the board members were women; and business men and professional workers accounted for more than 60 percent of the members.

THE ANNUAL WORKSHOP IN READING RESEARCH will be held at the time of the American Personnel and Guidance Associations meeting in Boston, April 9, 1963. The primary purpose of the workship is to encourage continued evaluation and research in learning as it is related to reading and communication.

THE NORTH ATLANTIC TREATY ORGANIZATION announces that a limited number of advanced research fellowships is offered for 1963-64 to candidates from member states (Belgium, Canada, Denmark, France, Federal Republic of Germany, Greece, Iceland, Italy, Luxembourg, Netherlands, Norway, Portugal, Turkey, United Kingdom, and the United States). Candidates will be selected on the basis of their special aptitude for and experience in carrying through a major project of research. In making selections, factors such as academic qualifications (the doctoral degree or its equivalent), professional experience, and publications will be taken into account. Further information on the NATO Science Fellowship Program may be obtained from the Fellowship Office, National Academy of Sciences-National Research Council, 2101 Constitution Avenue, Washington 25, D.C.

THREE INDEPENDENT DIMENSIONS SEEM TO BE most relevant in describing the academic environments of colleges. George Stern of Syracuse University and Robert Pace of the University of Southern California have been working for several years on the development of two instruments for studying the student and his needs in relation to his college environment and how it presses upon him. Using these instruments, they found the three dimensions to be: the intellectual climate, dependence versus autonomy, and emotional expression versus control.

Readers are invited to send items of research interest to John Schmid, Professor, College of Education, University of Arkansas, Fayetteville, Arkansas.

Among Those Present

RUTH CATES BAIRD is a teacher of English and reading at Oak Ridge, Tenn. She holds degrees from Carson Newman and Vanderbilt, and has done graduate work at George Peabody, Columbia, Duke, Northwestern, Chicago, Tennessee, and Harvard. She has appeared in *The Tennessee Teacher*, *The High School Journal*, and *The Journal of Developmental Reading*.

RICHARD H. BLOOMER is on the staff of the School of Education at the University of Connecticut. He holds two degrees from Columbia and a doctorate from Southern California. Dr. Bloomer has been associate professor and director of the educational clinic at the State University of New York College of Education at Geneseo, and senior scientist on the driver behavior research project and associate professorial lecturer at The George Washington University.

JAMES W. CREASER is with the Student Counseling Service at the University of Illinois, Chicago Undergraduate Division. He holds degrees from Michigan and Ohio State, and is currently working on a PhD at the Illinois Institute of Technology.

FREDERICK R. SMITH, assistant professor of education at Indiana, holds a BA from Duke, an M.Ed. from Boston, and a PhD from Michigan. He formerly taught in the secondary schools at Jackson, Mich.

RACHEL D. WILKINSON is an assistant professor in the Yeshiva University Graduate School of Education. She has served an the faculty of Hunter College and is a former dean of women. Her PhD is from NYU. She is sketched in Who's Who Among American Women, is a member of the New York City branch of the AAUW, and is a member of the alumnae board of the Harvard-Radcliffe program in business administration.

ROBERT H. KNAPP is professor of psychology at Wesleyan University and chairman of the department there. He holds two degrees from Oregon and a PhD from Harvard. During the last 10 years he has served as a consultant to the Ford Foundation, the Hazen Foundation, the Office of Naval Research, the U.S. Air Force, the Connecticut State Hospital, the Educational Testing Service, the Maxwell Graduate School of Citizenship and Public Affairs at Syracuse University, the Council on the Humanities at Princeton University, the National Science Foundation, and the Association of American Colleges. Besides some 40 journal articles, he is the co-author of *The Origins of American Scientists*, and *The Younger American Scholar*. LINDA A. EHLINGER is a 1962 graduate of Wellesley College.

JAMES V. MITCHELL, JR., a Chicago PhD, is associate professor of educational psychology at the University of Texas, has also been on the staffs of Miami and Purdue. He was recently the recipient of an award for an experimental design from the Television Bureau of Advertising.

ROBERT M. ROTH is director of psychological services and associate professor of psychology at Hampton Institute, and holds rank in the department of psychology at the George Washington University as well. He holds a BA from Indiana, an MA from CCNY, and a PhD from Texas, where he was formerly a Hogg Foundation fellow and clinical psychologist. He has previously appeared in a number of learned journals.

SIDNEY J. DRUMHELLER is assistant professor at the State University of New York College of Education at Brockport. His BA is from Denison and his EdD from Columbia. His experience includes four years as a secondary-school guidance director and five years as a college teacher.

Executive Editors

WILSON B. THIEDE, Professor of Education, Educational Administration, and Agricultural and Extension Education, The University of Wisconsin, Madison, Wisconsin. Chairman.
THEODORE HARRIS, Professor of Educational Psychology and Education, The University of Wisconsin, Madison, Wisconsin.
L. JOSEPH LINS, Professor and Coordinator of Institutional Studies, The University of Wisconsin, Madison, Wisconsin.
PHILIP LAMBERT, Professor of Educational Psychology and Educational Administration, The University of Wisconsin, Madison, Wisconsin.
JOHN SCHMID, Professor of Education and Chairman of the Department of Educational Foundations, University of Arkansas, Fayetteville, Arkansas.

Editorial Board

JACOB O. BACH, Head, Educational Research Bureau, Southern Illinois University, Carbondale, Illinois.
WILLIAM A. BROWNELL, Dean, School of Education, University of California, Berkeley, California.
ROBERT A. DAVIS, Professor of Educational Research, George Peabody College for Teachers, Nashville, Tennessee.
GUY D. McGRATH, Dean, College of Education, Arizona State University, Tempe, Arizona.
HAROLD E. MITZEL, Director, Office of Research and Evaluation, Division of Teacher Education, College of the City of New York, New York City.
WILLARD C. OLSON, Dean, College of Education, University of Michigan, Ann Arbor, Michigan.
ARTHUR E. TRAXLER, Executive Director, Educational Records Bureau, New York 32, New York.

Associate Editors

HAROLD M. ANDERSON, Associate Professor of Education, University of Colorado, Boulder, Colorado.
VERNON E. ANDERSON, Dean, College of Education, University of Maryland, College Park, Maryland.
RUDYARD, K. BENT, Professor of Education, University of Arkansas, Fayetteville, Arkansas.
EMMETT ALBERT BETTS, Research Professor, University of Miami, Coral Gables, Florida.
LAWRENCE P. BLUM, Assistant Director, Student Personnel Services, University of Wisconsin-Milwaukee, Milwaukee 3, Wisconsin.
J. MURRAY LEE, Chairman, Department of Elementary Education, Southern Illinois University, Carbondale, Ill.
D. WELTY LEFEVER, Professor of Education, University of Southern Cafifornia, Los Angeles, California.
S. E. TORSTEN LUND, Associate Professor of Education, University of California, Berkeley, California.
WILLIAM A. McCALL, Lecturer in Education, University of Miami, Coral Gables, Florida.
DELYTE W. MORRIS, President, Southern Illinois University, Carbondale, Illinois.
VALWORTH PLUMB, Chairman, Division of Education and Psychology, University of Minnesota (Duluth Branch), Duluth, Minnesota.
DAVID G. RYANS, Rand Corporation, Santa Monica, California.
LESTER B. SANDS, Professor of Education, University of California, Santa Barbara, Goleta, California.
MANDEL SHERMAN, President, and Consultant, Emotional Health Association, Los Angeles, California.
MILDRED B. STANTON, Consultant, Special Education, State Department of Education, Hartford 15, Conn.
W. W. THEISEN, Professor of Education, Marquette University, Miwaukee, Wisconsin.
M. R. TRABUE, Professor of Higher Education, University of Kentucky, Lexington, Kentucky.
J. WAYNE WRIGHTSTONE, Director, Bureau of Educational Research, Board of Education of City of New York, Brooklyn, New York.

Editors Emeriti

E. J. ASHBAUGH, Dean Emeritus, School of Education, Miami University, Oxford MR50, Ohio.
HARRY J. BAKER, 19050 Wiltshire, Lathrup Village, Michigan.
FOWLER D. BROOKS, Director Emeritus, Department of Psychology, DePauw University, Greencastle, Ind.
LEO J. BRUECKNER, Professor Emeritus, University of Minnesota, Minneapolis, Minnesota.
B. R. BUCKINGHAM, 33 Chatham Circle, Wellesley Hills, Massachusetts.
EDGAR A. DOLL, Director Emeritus, Devereux Schools, Devon, Pennsylvania.
HARL R. DOUGLASS, Director Emeritus, School of Education, University of Colorado, Boulder, Colorado.
E. DUNCAN GRIZZELL, Dean Emeritus, University of Pennsylvania, 640 Maxwelton Court, Lexington, Ky.
ERNEST HORN, Professor Emeritus of Education, State University of Iowa, W307 East Hall, Iowa City, Iowa.
EARL HUDELSON, Professor Emeritus of Education, West Virginia University, Morgantown, West Virginia.
A. M. JORDAN, Professor Emeritus, University of North Carolina, Chapel Hill, North Carolina.
ALFRED S. LEWERENZ, Director Emeritus of Evaluation and Research, 4107 Holly Knoll Drive, Los Angeles, California.
A. R. MEAD, Director Emeritus of Educational Research, University of Florida, 1719 N. W. 6th Ave., Gainesville, Florida.
M. J. NELSON, Dean Emeritus, Iowa State Teachers College, 6033 West Orange Drive, Glendale, Arizona.
JESSE B. SEARS, Professor Emeritus, Stanford University, Stanf
Department

Arvil S. Barr, Deceased, Editor and Publisher, 1928-1962

The Journal of Educational Research

Dedicated to the Scientific Study of Education

Volume 56 FEBRUARY 1963 Number 6

It Says in the JER ---------------------------------- 282

Summary of Investigations Relating to Reading, July 1, 1961 to June 30, 1962—Theodore L. Harris ---------------- 283

A Study of Experimental Methods for Teaching Emotionally Disturbed, Brain Damaged, Retarded Readers—Max Talmadge, Anthony Davids, Maurice W. Laufer -------- 311

A Comparison and Evaluation of Two Methods of Organization For the Teaching of Reading—Reginald Kierstead ---- 317

The Relationship Between Intelligence Quotients and Gain in Reading Achievement with Arithmetic Reasoning, Social Studies and Science—Carrie M. Scott ---------------- 322

The Effectiveness of Workbooks in the Elementary School Reading Program—Robert L. Docter ----------------- 327

Simplified Handwriting for Today—Gertrude Hildreth ------ 330

Field News—John Schmid ---------------------------- 334

Among Those Present ------------------------------- 336

The Journal of Educational Research is published at Madison, Wisconsin, monthly 10 times a year (th
May-June and July-August issues being combined). The subscription price is $7.50 a year, plus $
postage for all subscriptions outside the continental United States. Single copies, $1. Second clas
postage paid at Madison, Wisconsin, Copyright 1963 by Dembar Educational Research Services, Inc
Address all correspondence care of DERS, Box 1605, Madison 1, Wisconsin.

Published by **DEMBAR EDUCATIONAL RESEARCH SERVICES**, Inc. WALTER FRAUTSCHI, President
Prof. WILSON B. THIEDE, Vice President and Publisher. Prof. CLARENCE A. SCHOENFELD, Assistan
to the Publisher. ARNOLD CAUCUTT, Treasurer and Business Manager. JEAN WILLIAMS POTTER
Editorial Services.

It Says in The JER...

FROM BOSTON'S EMMA PENDLETON BRADLEY HOSPITAL COMES A UNIQUE APPROACH TO READING retardation involving kinesthetic and auditory aids applied to children with central nervous system involvement who are also emotionally disturbed. Talmadge, Davids, and Laufer suggest the importance of individual attention and stimulating techniques.

WHEN GRADE LINES ARE ENTIRELY REMOVED FROM THE READING PROGRAM DO YOU GET MORE PROGRESS than when ability is locked within each grade? Kierstead reports "no significant differences" made in vocabulary skills and reading comprehension when the groups are equated on the basis of appropriate intelligence and basic-skills tests. Teachers seem to prefer the ungraded structure, but students in "slow" sections feel a social stigma.

IMPROVED READING ABILITY MAY LEAD TO IMPROVEMENT IN ARITHMETIC REASONING, SOCIAL STUDies, and science, but the degree of correlation is less than many educators realize, Scott reports from an Oregon study.

DO WORKBOOKS HELP TEACH READING IN THE SIX ELEMENTARY GRADES, OR DO THEY REPRESENT mere "busy-work?" In California tests by Docter, non-workbook classes did better at the first grade level, and the workbook classes did better in grades two, three and four. No significant differences showed up in grades five and six.

GERTRUDE HILDRETH MAKES A PLEA FOR WIDER ADOPTION OF AN ADVANCED TYPE OF MANUSCRIPT handwriting, known as "italic script," in place of block writing or cursive style.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 6, February 1963)

Summary of Investigations Relating to Reading: July 1, 1961 to June 30, 1962*

THEODORE L. HARRIS
University of Wisconsin

THE POLICIES which guided the preparation of the previous Summary (46) have been followed in the current Summary: 1) to reflect the broad scope of research in the sociology, psychology, physiology, and teaching of reading, 2) to include empirical investigations only, 3) to consult primary sources insofar as possible, and 4) to identify experimental designs and methods of analysis as well as the research findings. This has been done with the cognizance that while many selective bibliographies appear rather regularly in other journals,** the distinctive function of the Summary which has appeared for many years in the *Journal of Educational Research* will be best served by the policies stated.

A number of studies, including a group reported in or based upon *Dissertation Abstracts*, and others reported in the proceedings of various conferences, were judged too incomplete in presentation for inclusion in the Summary. It is hoped that such studies will be published in more complete form.

The Sociology of Reading

1. Environmental Influences Upon Reading

Social Mobility and Reading. When children of similar academic ability are compared, does change of schools adversely affect reading performance upon standardized reading tests? In order to confirm or disprove the negative answer to this question found in a 1952 study in Cincinnati public schools, Bollenbacher (14) analyzed the number of individual changes of schools made by 5578 sixth grade pupils in three socio-economic areas of the same city in 1960. The earlier finding was in general confirmed. Pupils in schools in suburban, middle, and "basin," or downtown, areas moved on the average 1.30 times during a five and one-half year period, but this figure reflected largely the approximate one-third of the class that moved more than once. Over three-fourths of this mobile group were in the non-suburban areas. The study thus indicated that mobility was associated with the less privileged areas of the city, as were the lower median levels of reading achievement and intellectual ability, but that mobility *per se* did not appear to be associated with the expected level of reading achievement of the large group of pupils studied.

Effect of Television Viewing. Belson (6) described the effect of television viewing upon newspaper and magazine reading habits of a representative sample of 1,157 adults in London, 288 of whom were television viewers and 869 of whom were non-viewers. The sample was considered similar in such factors as age, sex, occupation, job level, marital status, and age of leaving school. Results were described in terms of types of newspapers and magazines read, amount read and topics read about in 53 publications. It was concluded that the total effect of television viewing upon the reading of newspapers and magazines was small, though positive effects were noted in increased reading of newspapers, especially popular ones, and in increased reading about items actually dealt with on television programs. Hess and Goldman (47) noted from interviews with 99 mothers in a metropolitan area that 78% of them agreed with this statement concerning the effect of television on their children: "The more a child watches television, the less he reads." Sorelle and Walker (94), in a comparative study of the same population of 2000 students first in junior and then senior high school, found that a larger percentage (40%) read less than before television as junior high school students than as senior high school students (33%).

---

* With the assistance of Mrs. Barbara Lewinson and Mrs. Carol Lindeman.

** See other bibliographies and reviews of research in reading in *Elementary English*, *Journal of Developmental Reading*, *The Reading Teacher*, *Review of Educational Research*, and the *Proceedings* of the International Reading Association and The National Reading Conference.

2. Characteristics of Reading Materials

Characteristics of Children's Readers, 1800-1950. DeCharms and Moeller (25) analyzed trends in achievement and affiliation imagery, and of moral teaching, in children's readers from 1800 to 1950. A systematic sample was taken of four readers, designed for fourth grade use, at each 20-year interval. Every third page was scored for a total of 2,375 pages. Among the findings of interest, it was found 1) that achievement imagery increased from 1800-1890 and then declined, 2) that affiliation imagery tended to increase to 1930, then dropped sharply, and 3) that moral teaching decreased, very sharp drops being noted after 1850 and after 1910. These and other findings were discussed in relation to factors influencing cultural change.

Content Analyses of Newspapers. The degree of emphasis upon editorial content was the focus of two recent studies. Peterson and Thorp (83) made a comparative analysis of the proportion of editorial content in 215 current non-daily Iowa newspapers in November, 1960, and that in an earlier 1930 survey. They found that only 41 percent of the current newspapers had editorial pages compared to the 63 percent found in 1930. Furthermore, only 27 percent of current weeklies with circulations under 2500 had editorial pages. The authors concluded that the personal columns which have replaced the editorial page are a weak substitute for the latter.

An analysis of space devoted to editorials in current daily newspapers, is, however, perhaps more encouraging. Stempel (99), in an analysis of the content of the four largest and four smallest Michigan daily newspapers, noted that the small-town daily tended to place a little more emphasis upon news and editorial comment. The study was based upon a one week's issue, exclusive of Sunday. Not all small-town or all big-town dailies were found to be alike with respect to the 11 categories of content analysis used. The results of linkage analysis, including clusters which characterize the total group of newspapers examined, are detailed in the article.

Analysis of the Radical Rightist Press. Wilcox (110) studied a sample of publications of the radical rightist press, excluding general newspapers and religious tracts, drawn from the First National Directory of "Rightist" Groups, Publications and Some Individuals in the United States and Some Foreign Countries in order to identify thematic patterns and the position of these publications on the political opinion spectrum. In the extensive phase of the study, 42 publications represented by one to five issues each were selected for analysis from an original group of 110 publications. Each article was placed upon a five-point scale with respect to two major dimensions, commitment to conspiracy and non-rationality, and upon a four-point scale with respect to two minor dimensions, the black-white, evil-good dichotomy and that of negative, non-constructive criticism. The results of the scaling procedure showed that this technique adequately measured the radical rightist press in terms of the concentration of scale values at the upper or high end of the continuum in the various dimensions. It was less efficient at the lower end of the continuum which was marked by a slight rise in scale values of a more neutral type. Examination of the thematic characteristics of publications at the low end of the scale indicated that the latter did not fit the dictionary definition of "conservative." The author likewise made an intensive analysis of two publications using a "recurring value word count" technique and drew certain inferences regarding the rightist press based upon the results of the total study. The article should be examined for details of the scaling categories and techniques of analysis used.

Stylistic Attributes of Sensationalism. In a study which represents an extension of an earlier one in 1960, Tannenbaum and Lynch (103) attempted to identify the stylistic attributes of passages read by 61 adults to relate these to the degree of sensationalism (Sendex score) of the passages. Subjects read 24 passages from a college level reading test and rated each passage against a set of semantic differential scales representing the three main dimensions of sensationalism. Each subject also judged the concept "sensational news" against the same set of scales. Correlation and factor analysis techniques were used to identify from among 37 separate stylistic indices, three stylistic clusters which appeared to be most parsimonious as predictor variables. The following clusters represented 60 percent of the Sendex score: 1) readability, as measured by the Flesch score combining number of syllables per 100 words and words per sentence, 2) punctuation usage, as indicated by the amount of internal punctuation per sentence, and 3) degree of modification, as shown by the ratio of the total number of adverbs and adjectives to the total number of nouns and verbs. Some validation of the original judgmental measures was secured by encoding a set of news stories to represent three different levels of sensationalism on each of three topics. Analysis of these revealed almost the same values obtained as in the general reading passages. The authors cautioned that there is no evidence for the generality of these findings nor for the conclusion that the identified clusters represent the basic underlying message features.

3. Reading Habits

Newspaper Readership. Lyle (61) conducted a questionnaire survey of the newspaper reading habits of 674 sixth graders and 700 tenth graders, a population which included almost all students in these grades in five Rocky Mountain towns. On the assumption that the reading of entertainment sections of the newspaper constituted "immediate" reward, while reading "hard" newspaper sections such as news and editorials constituted "delayed" reward, Lyle examined several variables and concluded that typologies do exist and change with age. Significant relationships were found between the selection of newspaper content and 1) media use in general in grade 10; 2) frequency of newspaper reading (tenth graders read more in each category); 3) teacher's estimates of intelligence; 4) socio-economic status according to father's education and occupation; 5) conflict with parental aspirations in grade 10; 6) discussion of news (greater in grade 10); and 7) college aspirations (greater for hard news readers at both grade levels).

The readership of the editorial page, as well as the space devoted to it, continues to be an object of research interest. Baker and MacDonald (4) used data from *The Continuing Study of Newspaper Reading, 1939-1953*, to determine the relationship between selected editorial variables, including length, and the number of persons attending to them, with respect to the readership of a country-wide sample of 130 different daily newspapers. The basic analysis was made upon the classification of three lengths of editorials and three percentages of readership, sub-divided by sex, plus further sub-categories of content and readership characteristics involving a total of 11 variables. Among the results it was noted 1) that the average editorial readership of 22.3% compared very favorably with averages for other types of content; 2) that a lead position of an editorial tended to attract readers and 3) that editorial readers in general in this study tended to read something if they looked at an editorial page, read editorials regardless of length or difficulty, and read all or most offerings on the editorial page.

Readership and Political Preference. Stempel (100) conducted interviews with 205 college students to determine if they read only that material agreeing with their political views, in this case in relation to two articles circulated about the views of candidates for campus offices and to the two platform statements of the respective candidates. The investigator found 1) that the majority of the students read both the articles and the platform statements, especially the latter; 2) that a smaller but substantial percent read all of the articles and platforms (44-58%); 3) that a higher percent (49%) read equal amounts about both candidates than read more about candidates of their choice (31%); and 4) that students read more about candidates of their choice than about the non-preferred candidate. The investigator concluded that the readership of political content in this group is higher on this occasion than one would have reason to believe for daily newspaper reading.

4. Psycho-Social Effects of Reading

Early Reading to Children. Inasmuch as reading and language development are closely related, Irwin (50) conducted an experimental study during a 17-month period of the effects of systematic reading of stories to children of between 13 and 30 months upon the amount of phonetic production. An experimental group of 24 children and a control group of 10 children were drawn from working class homes, and only mothers of the experimental group read stories to the children from 15 to 20 minutes daily from books supplied by the experimenter. Spontaneous speech was recorded and analyzed once every two months. Little difference in mean phoneme frequency scores appeared during the first six months, but after this point, with one exception during the 21-22 month interval, significant differences favoring the experimental group were found to be at or beyond the five percent level of confidence. This finding is pertinent to the oft-claimed values of reading to children during their early years.

Headline Motivation. Felstehausen (32) conducted interviews with 200 male and 201 female readers of specially prepared agricultural news articles to study the effect upon readership scores when the same article was presented under headlines written to provide different motivational appeal. No consistent preference was found for either type of motivation. The subjects represented a sampling of a state-wide population.

Reader Attitudes and Statements of Opinion. In extensions of a previously reported study of the interpretation of statements of opinion of college students toward fraternities, Manis (70)(71) conducted two additional studies following the same general methodological structure. Attitudes toward college fraternities were explored by ratings on six semantic differential scales for 66 and 65 male undergraduates, respectively, in the two studies to derive anti-fraternity, pro-fraternity, and neutral groups. In the first study it was found that in response to ambiguous statements about fraternities, secured by deleting every other word, subjects tended to interpret statements with which they had moderate or no disagreement more toward their own bias. In the second study, pro-fraternity subjects tended to assume that high prestige statements would be more pro-fraternity than those of low prestige, while anti-fraternity subjects showed some tendency to reverse this assumption.

Byline Placement and Attitudes. Greenberg and Tannenbaum (39) used semantic differential analysis techniques to determine pre- and post-attitudinal effects of the presence, placement or absence of bylines reflecting high prestige sources on the reading of three persuasively presented topics by 216 university undergraduates. Four semantic differential scales were employed: good-bad, valuable-worthless, wise-foolish, and important-unimportant. Each subject read three persuasive communications on different topics under four different byline conditions: byline at beginning, near the middle, near end, or no byline. Factor change scores were computed which showed that there were no significant differences between the beginning and middle placement of the byline, but that both of these conditions produced greater attitudinal effects than did the absence of a byline or its placement near the end of the message. The authors concluded that the critical factor was not where the byline was located but when the reader perceived it, early identification producing greater changes in attitude.

Media Meaning and Reader Interest. In exploring possible media message effects upon reader interest in unfamiliar foreign news topics, McNelly (68) made three basic assumptions: 1) that meaning intensity, as measured by semantic differential scales for potency, activity and three types of evaluative judgments, will increase with the reading of the topic; 2) that the increase in meaning intensity will be accompanied by an increase in expressed interest in the topic similarly measured by appropriate scales; and 3) that there will be a corresponding increase in the probability of selective readership about the topic, as measured by a forced-choice test of alternative headlines. Factual news stories on eight different foreign news topics, each unfamiliar, were read by 104 high school and 126 college students under randomized conditions of distribution of the topics.The results of the study of general confirmed the three basic assumptions of the study, the majority of the findings for the eight topics being related to each postulate at or beyond the five percent level of confidence. The author also found a rank order correlation of .95 between the headline test and meaning intensity and one of .83 between the headline test and expressed interest.

The Psychology of Reading

1. Psychological Correlates of Reading

Determinants of Comprehension. Vernon (106) reopened the long standing controversy of essay vs. multiple-choice techniques for measuring reading comprehension and that of general vs. specific factors involved in reading comprehension by constructing or adapting a series of 11 tests and administering them to 108 male English college students and 75 American college students of both sexes. In the appraisal of the "relative variance of certain test-content, method, and error components," pairs of tests calling for multiple-choice vs. "creative" (written in own words) responses were devised to measure vocabulary, comprehension with and without re-reading, plus other tests of sentence completion, reading graphs, etc. Two forms of each test were prepared and administered alternately to half of each group in sessions of a week of more apart. Mean scores, standard deviations, and inter-session gains were reported, and test intercorrelations calculated and studied by centroid factor analysis. The specific results of the study were related to eight hypotheses reflecting the basic question posed by the author: "While admitting that fluency and legibility in essay writing unduly affects the students' performance in any European examination, he wonders whether new-type test sophistication does not equally distort assessments of abilities in the American setting." Among the major findings of the study, it was noted that 1) "Assessments of the understanding of complex concepts... are affected... by many factors arising from the methods of testing, and the subject's facility in handling it. Certain weaknesses in current new-type achievement tests, such as their poor differential predictive capacity, may result from neglect of such method components."; 2) the more highly selected British group was superior on vocabulary but not comprehension tests, but its significant improvement with practice on the latter "illustrates the importance of facility or sophistication at such tests."; 3) higher correlations were found among vocabulary and reading comprehension scores within tests than between sessions, especially in the American group; 4) a comprehension test involving retention with no opportunity for re-reading "appeared to measure a somewhat different ability from the conventional immediate-comprehension test, and it was found to be considerably more valid than the latter in the prediction of academic achievement;" 5) "The writing of responses to vocabulary and reading questions by students in their own words (creative type) did not, as had been hypothesized, involve a different ability from the objective or multiple-choice type of response;" and 6) "Centroid analyses revealed a strong Comprehension factor, orthogonal to the Vocabulary factor, among both groups in the reading tests." Its validity for predicting academic achievement was positive but low for American students and slightly negative for British students. The author accepted these findings as evidence that tests of different levels, as vocabulary and reading comprehension, involve partially distinguishable abilities, more in line with Davis' position with regard to the specificity of reading skills than with Thurstone's more unitary view.

Personality and Reading. Walters, Loan and Crofts (107) sought to test certain inferences drawn from a psychoanalytic theory of reading disability; specifically, that retarded readers would show greater hesitancy in looking at a sexual object than would good readers, that this "avoidance" would also be evident in perceptual tasks of a neutral type, and that retarded readers would display hostility toward the same-sex parent with attendant evidence of aggression and "faulty identification mechanisms." The subjects were 18 good, 20 average, and 20 poor readers from grades three through six, the latter group having reading ages one year or more below their mental ages. The testing procedure included 1) a multiple choice box with three doors behind which could be hidden a Dutch doll, a nude male doll, and a clothed male doll on successive trials, so arranged that the elapsed time in opening the doors after a given signal could be measured; 2) a parent preference test by telling boy he could go on a trip with one parent, then flipping a coin and asking boy to call out preference on a series of trials; 3) a projective picture-story test; 4) the Steer-Allen Figure-Ground Confusion Test; and 5) the Steer-Beatty Closure-Threshold Test. It was found that the retarded readers did more poorly on the two form perception tasks and that they were significantly slower in opening the compartment door to look for a nude male doll. They likewise chose the father less frequently in the parent preference task. Fantasy data from the picture-story test, however, failed to support either the fear of looking or hostility hypotheses.

Otto and Lasswell (82), observing that personality attributes associated with the retention in adolescence of a secondary beginning stroke in handwriting were often found in poor readers, analyzed differences in the presence of such a stroke among 25 good and 25 poor readers of normal intelligence in grades 7 and 8. The number of secondary strokes found in a sample of ten words beginning with the small letters u, v, w, and y was determined from the subjects' themes and spelling exercises. Chi-square analysis showed that poor readers used a greater number of secondary beginning strokes than did good readers, a finding significant at the one percent level of confidence, while the seventh graders used more such strokes than did eighth graders. The study is of interest because of the possible diagnostic significance of a characteristic interpreted as a symptom of immature personality development in handwriting research as that by Linton, Epstein and Hartford and the presence of similar personality characteristics in poor adolescent readers.

Ramsey (84) administered a 20-item questionnaire to 61 poor readers and 81 good readers in grades 4, 5, and 6 to study their reading interests habits and self-concepts as readers. Responses to two items relating to their self-concept are of interest. When asked to rate themselves as readers, good readers responded thus: good or very good, 50%; average, 47%; poor, 3%. Poor readers responded: good or very good, 18%; average, 49%; poor, 3%. Poor readers responded: good or very good, 18%; average, 49%, poor, 33%. Since the poor readers were at least one year below grade in reading achievement, the majority of their estimates of their reading ability were hardly realistic. When asked to express their felt needs in reading, good and poor readers differed considerably in the percent of needs expressed in all but one category (percent for poor readers is given first): word analysis, 23.5%, 7.5%; increased speed, 36%, 27.5%; improved comprehension, 25%, 26%; and increased vocabulary, 15.5%, 39%. The fact that increased speed and vocabulary were the most and least frequently expressed needs, respectively, by poor readers is in sharp contrast to those expressed by good readers for the same categories.

Attitudes and Reading. Groff (41) explored the relationship between the performance of an "average" population of 142 boys and 163 girls in grades five and six on an experimental critical reading test and their expressed attitudes on modified existing attitude scales toward reading, school and four types of reading content: stories about boys' sports, girl's mild adventure, airplanes or flying, and manners or social relations. Correlations between attitudinal and reading test scores for the four types of content and other personal variables, including sex, were computed. Major results of the study showed 1) that correlations between attitudes toward content types of materials and critical reading test scores on four types of content were all significant at the one percent level of confidence for the total group; 2) that this relationship was more consistent for girls than for boys; 3) that a similar positive relationship "appeared likely" between attitudes toward reading and critical reading scores; and 4) that the relationship between general reading ability and attitudes toward four types of content was negligible for the total group, but significant by sex.

Carner (18) used two questionnaire forms to explore the attitudes of 111 fifth and sixth grade pupils with respect to their before and after reaction to televised reading instruction three times per week. The lowest 27% of the scores, representing a negative attitude, were analyzed in relation to reading achievement. As measured by the questionnaire, change significant at the one percent level of confidence occurred for the group as a whole and for superior and average readers, in the direction of a positive attitude, but not for below average readers.

Emotional feelings toward words were also examined. Bloomer (11) investigated the relationship between word length and spelling difficulty, and their emotional intensity and tone judged by college juniors. Indices of reading and spelling difficulty were obtained from ratings by Gates and by Greene for 149 words selected at random from the New Iowa Spelling Scale. The emotional tone, positive or negative, was rated by 45 subjects for the words randomly arranged in a larger 500 word list, while 47 subjects rated the 149 words as strong or weak as an indication of emotional intensity. In each case, the sum of the judgments of the group constituted the score for a word. Word frequency was determined from the Rinsland study. It was found in general that while frequency of occurrence and word length were significantly related to the word difficulty, there was no evidence that the latter was related to the ratings of emotional tone or intensity for this sample of common school words.

Reading Interests. To what extent are the interest categories in pre-primers and primers similar to those represented in the free choice of library books by first graders? Smith (92), using 23 interest categories, classified the centers of interest of 37 pre-primers and primers and compared these with those of the library choices of 110 first grade pupils. The first five free-library choices of the pupils were, in order: humor-fantasy, real animals, nature, holidays and birthdays, and fairy tales. The five most frequent interest centers in 859 selections examined in pre-primers and primers were, in order: children-parents, real animals, toys, play-games, and humor-fantasy. Eight areas of interest shown by the children's library choices which were not reflected in the readers at this level were: adventure-mystery, fairy tales, Indians, space, astronomy, other countries, wild west cowboys, and long-ago history. On the other hand, the categories of clothes and food found in the readers were not reflected in any library choices of the pupils. Smith concluded that the interests of children are considerably broader than those represented in their beginning reading books, and that the latter do not satisfy children's reading interests as shown by their choice of library books. The study reinforces the need for a wide range of materials for children's reading.

Stanchfield (97) employed interview and rating scale techniques to study the reading interests of 51 eighth grade boys selected by a stratified random sampling procedure from a large metropolitan population. Three sub-groups of 17 boys each were classified as good, average, or poor readers on the California Reading Test, and each subject met additional criteria of being within a 90-120 IQ range, having good physical and mental health, and being American-born. In interviews which lasted over an hour per subject, general attitudes toward reading and factors affecting reading interests were explored, including verbal rating on four-point scales of 50 topics into interest categories and of 20 qualities of writing and style according to characteristics of reading interests. The results of mean ratings on the two scales revealed no significant differences among the three reading group levels. Characteristic reactions of frustration accompanied by a sincere desire to do well were noted among poor readers, and the importance of better reading material matched to reading interests of this group was stressed.

Martin (72) asked each of 250 sixth graders to name the three best books they had read. They were instructed to select books which they found so interesting it was difficult to put them down before they had finished. The fifty books receiving the most frequent mention are listed in the article.

Intellectual and Achievement Correlates. In an extension of a previously reported 1961 study, Durkin (27) reported on the further reading progress through the second grade of a group of 49 children who had learned to read before entering school in relation to certain characteristics accompanying their early introduction to reading. As a group, these children made steady progress in reading achievement on standardized tests during six successive testing periods, attaining a median grade level reading score of 4.9 by the end of second grade. Data are also cited in relation to reading progress and intellectual level which show 1) that the modal age of beginning instruction in the home was four years; 2) that the frequency of instruction was either very often or rather often in the great majority of cases rather than intermittent; 3) that the intent of the parents was not deliberate in the great majority of instances; and 4)that the instructor was usually either a parent or a combination of parent and sibling. The two studies serve a valuable function in exploring the conditions associated with learning to read before formal schooling.

McDonnell (64) correlated three group intelligence tests with scores on the primary battery of the California Achievement Tests for 178 third grade Canadian children with IQ's of 120 or more on the California Test of Mental Maturity. The intelligence tests used were the Detroit Beginning First-Grade Intelligence Test, the California Test of Mental Maturity – Short Form, and the Raven's Progressive Matrices. Intercorrelations among intelligence tests were low, ranging from .10 to .19, and among these and California Achievement Tests moderately low, ranging from .08 to .36. It was found that the Detroit and the Progressive Matrices were the best predictors of total achievement and achievement in reading, language and arithmetic on the California Achievement Test for these more able pupils.

Curry and Hughes (24) analyzed certain relationships between anticipated and measured achievement in reading vocabulary and comprehension among 640 students in grade 10, using the CTMM and the advanced California Achievement Test as the bases of measurement. Students were classified by IQ as above average (IQ 110 +), average (IQ 90-109), and below average (IQ 89 -) in the analysis. Differential results were found for the group as a whole and for groups classified according to intellectual ability by comparing their mental or anticipated grade placement on the CTMM with their measured grade placement in reading vocabulary and comprehension. Correlations for the whole group tended to be substantially higher than for the sub-groups. In general it was found 1) that the above average group tended to achieve above grade level but slightly below their anticipated grade placement; 2) that the average group tended to achieve above grade placement and above anticipated grade placement; and 3) that the below average group tended to achieve below grade placement but above anticipated grade placement. It was also noted that the general level of achievement in reading comprehension was higher than that for vocabulary, and that the relationship between anticipated and measured achievement for the group as a whole was relatively high in vocabulary but moderately low in reading comprehension.

The Gates Reading Survey was used by Cleland and Toussaint (19) as a criterion measure in studying the interrelationships among reading, listening, arithmetic computation and scholastic ability for 172 pupils in grades 4, 5, and 6. Pertinent findings related to reading were 1) the STEP 4a Listening test was more highly correlated with the criterion, .67, than was the Durrell-Sullivan Reading Capacity Test, .58, another listening test; 2) the STEP Listening and the Primary Mental Abilities Tests together yielded a multiple r of .76; and 3) the addition of the American School Arithmetic Computation Test to the latter two variables increased the multiple r only slightly to .79. All intercorrelations among these tests and for the Stanford-Binet were positive and were reported significant at the one percent level of confidence. The Stanford-Binet and the Primary Mental Abilities tests showed substantially the same relationship to the Gates Reading Survey, .61 and .59, respectively. The authors concluded that the relationship demonstrated between listening and reading justifies greater attention to listening activities in the instructional program at the intermediate level, a problem to which Lubershane (60) addressed himself.

Bentall (7) investigated the relationship between 1) reading level, 2) intelligence, 3) achievement in mathematics, English, science and social studies, and 4) related course marks for 830 ninth grade and 720 eleventh grade randomly selected students. Reading achievement was assessed on the advanced Stanford Achievement Test and upon the sub-tests of ability to interpret social studies, science and literary materials on the Iowa Tests of Educational Development. Data were analyzed according to mean group scores for degree of reading retardation below grade level (two or more years, one year, and no retardation) and by intercorrelations among the variables. Among the results of the study it was found 1) that the percentage of successful students increased with reading proficiency; 2) that students retarded by more than two years in reading at the ninth grade level were more handicapped in mathematics and science than in English and social studies; 3) that groups of science and mathematics students, particularly the ninth grade science group, were superior to English and social studies groups in IQ and reading proficiency; and 4) that correlations between teachers' marks, intelligence, and reading and subject-matter achievement were low, indicating that teacher's judgments apparently were influenced by other factors.

Keshian (56) examined differences between reading and mental ages among 406 fifth grade pupils selected by stratified random sampling from communities representing three socio-economic levels, with further analysis of 72 successful readers from this population. Only 11% of the total population was found to have reading ages on the Stanford Reading Test below mental ages on the Otis. For the selected sample of successful readers, socio-economic status was not found to be a differentiating factor, while other factors commonly found to be associated with successful reading were noted.

Correlations between reading achievement, scholastic aptitude and other aspects of school achievement may be found in the Educational Records Bulletin, Nos. 79 (28) and 80 (29).

Under- and Overachievement. Mingoia (74) conducted three studies with reference to underachievement and overachievement in reading to identify associated pupil and environmental characteristics among a population of over 3000 pupils in grades 4, 5, and 6. In one study involving 128 pupils with discrepancies of one year or more between the Durrell-Sullivan Reading Capacity Test and the Reading Achievement Test, such characteristics as "over-pressure," an "emotional block," and to a slight degree, bilingualism, were found associated with underachievement. A second study examined socio-economic conditions in relation to reading over- and underachievement among 3000 sixth graders, using the Kuhlman-Anderson and the Stanford Reading Test as identifying instruments. Overachievement (reading age one year or more above mental age) was rare in a low-socio-economic neighborhood with mean IQ's in the 88-92 range; overachievement approximated underachievement for neighborhoods with mean IQ's in the 98-103 range; and overachievement was frequent in privileged neighborhoods with a mean IQ range of 108-112. A third study examined the activities and attitudes of slight over- and underachievers, 30 each in the IQ range 105-115 in sixth grade in a privileged neighborhood. No differences were found on an attitude questionnaire toward reading teachers, classmates and out-of-school activities between the groups. The author concluded that no common characteristic distinguishes all underachievers, but cited several characteristics associated with extreme underachievement based upon the three studies.

Curiosity. Maw and Maw (73) tested the assumption that children with a high curiosity rating will tend to be more alert and actively seek understanding than will those rated low in curiosity. Curiosity ratings were secured by combining teacher, peer and self judgments of curiosity in two studies of fifth grade children, the pilot study including 102 children with a mean verbal IQ of 104 and the second, 458 children with a mean verbal IQ of 115. Comparison of the performance of children of high and low curiosity ratings upon a specially prepared "Foolish Sayings Test" in the pilot study yielded "t" test differences in the means favoring the high curiosity group at or beyond the one percent level of confidence. A similar comparison in the second study produced differences only at the five percent level of confidence. These findings led the authors to conclude that while children of a high curiosity level tend to sense the meaning of sentences more accurately than do those of low curiosity but of equal intelligence, the advantage of the high curiosity children is less at the upper IQ levels.

Language and Speech. Correlation studies of writing factors, awareness of sentence structure and grammar, speech development, and articulation errors in relation to reading also appeared.

Wyatt (112) examined the relationship between amount of reading and six language facets (use of subordination, spelling, capitalization, punctuation, vocabulary, and average sentence length) for within-group comparisons of 65 sixth grade children classified by scholastic ability: 21 with 126-140 IQ's, 21 with 111-115 IQ's, and 23 with 95-110 IQ's. Each child wrote four compositions, two on vicarious experiences and two on real experiences. Amount of reading was determined by a questionnaire and library records for the previous three years. No conclusive evidence was found that amount of reading within these groups was significantly related to the writing factors studied, although positive rank-order correlations at the five percent level of significance were obtained between amount of reading and the use of more subordination and longer sentences among the very bright, and between amount of reading and spelling success among the average group.

O'Donnell (80) prepared a test to measure awareness of sentence structure and studied the performance of 101 high school seniors upon this and upon the Cooperative Test of Reading Comprehension and the Iowa Grammar Information Test. The structural relationships reflected in the test were: 1) subject and predicate verb, 2) predicate verb and complement, 3) modifier and element modified, 4) coordinate elements, and 5) pronoun and antecedent. Test items consisted of a sentence pattern followed by three options in which nonsense words were used, the structural pattern of one option corresponding to the sentence pattern. Moderate correlations were secured betwen level of comprehension and awareness of structure (.44) and grammer (.46), and between vocabulary and awareness of structure (.46), but a correlation of .90 was found between vocabulary and awareness of grammar. The implications of these findings were discussed, and the point made that the good reader must consciously or unconsciously have an awareness of the basic structural relationships of words in sentences.

Sampson (87) used test data, clinical observations, interviews, and teaching ratings to study 48 eight year old children with particular reference to their speech and reading development. The correlations between reading comprehension at eight years and earlier speech results were as follows: with speech at two and one-half years, .69; with vocabulary at age five, .67, and with language at age five, .58. Other findings based upon clinical evidence and the use of rating scales led the author to conclude that in this study the relation of reading skill to the mechanics of speech was not clear, that good reading was not incompatible with maladjustment, and that underfunctioning in reading had an emotional and attitudinal component not necessarily associated with defective speech and language development. This British study should be consulted for further particulars.

Weaver, Furbee and Everhart (109) employed the technique of asking 638 first grade children to label a series of pictures designed to elicit the different consonant sounds in a clinical study of articulatory errors in relation to scores on the Gates Reading Readiness Test. Articulatory errors classified as substitutions, omissions, distortions, and additions were found to bear an inverse relationship to reading readiness scores; i.e., mean reading readiness scores tended to decrease as articulation errors increased, the correlation between the two measures being -.20 and reported significant at the one percent level. Chronological age did not appear to bear any relationship to articulatory competency in the group of children studied.

2. Developmental Factors in Reading

Children's Oral Language Patterns. The application of normative techniques of linguistic analysis to the study of the nature and development of children's language patterns was reported in an extensive monograph by Strickland (102). In this study the oral language patterns of a random sampling of 575 pupils in grades 1 through 6 were analyzed in relation to language patterns in basal reading textbooks and to other within-subject variables. A major hypothesis of the study was that "a study of children's speech, its structure and its patterns of arrangement and flow, may offer suggestions for the construction of better reading textbooks for beginners and possibly for older children as well." The basic method of the study involved the recording of children's spontaneous speech and the use of a method of linguistic analysis which permitted the identification and classification of phonological units with respect to: 1) the syntactic structure of sentences, 2) the frequency of occurrence of certain syntactic patterns, 3) the amount and kinds of subordination, 4) sentence length, and 5) the flow of language. A total of 14,375 phonological units or sentences were analyzed with the aid of a specially devised linguistic scheme which permitted the identification of "fixed slots" and "movables" in the sentence pattern and the types of subordination used. The same type of analysis was also applied to the language of basal reading texts. The relationship of oral language patterns was examined for several pupil and environmental characteristics, at successive grade levels, and the relationship of these patterns to silent reading comprehension, oral reading interpretation, and listening comprehension studied at the sixth grade level. Many specific developmental findings were reported in the study in addition to such general conclusions that 1) children learn the basic structures of their language at an early age;2) their oral language is far more advanced and varied than is that of typical basal readers; 3) children may need help "to recognize and understand the entire phonemic scheme;" and 4) "no evidence is available regarding the relationship which may exist between children's use of basic structural patterns and the degree of grammaticalness of their speech."

In particular, the following generalizations were developed from the study of representative basal reading series: "1. The basic subject, verb object pattern was the only pattern to appear in the samples of practically all of the books. 2. The patterns which appeared in the sample differed from book to book within a series as well as from series to series. 3. Patterns of sentence structure appeared to be introduced at random in a rather haphazard manner. 4. A pattern of structure, once introduced, seemed not to be followed by any sort of repetition or effort at mastery. 5. There appeared to be no scheme for the development of control over sentence structure which paralleled the generally accepted scheme for the development of control over vocabulary." (pp. 103-4). An important implication of this study is thus that the nature, variety and control of sentence structure in materials for teaching reading may well need reevaluation insofar as some congruence between the spontaneous speech patterns of children and the materials with which they learn to read is desirable.

Vocabulary and Comprehension. Gates (36) reported upon the relative ease and success which 412 pupils experienced in grades 2 and 3 in deriving the correct meaning of new words not yet encountered in their basal readers when these words were presented in multiple-choice completion contexts which provided only minimal context clues. In the earlier study of 310 third grade pupils, the children did substantially as well in identifying the correct meaning of the new fourth grade words as the familiar third grade ones. Similar results were secured in a second study of 47 second grade and 55 third grade pupils with respect to their ability to identify the correct meaning of new third and fourth grades ones. The author concluded that the effects of extensive reading and functioning word recognition skills are commonly disregarded in estimating the reading vocabularies of children, and noted that "The fact that a word is 'new' in the basal materials is an unreliable indication of its difficulty at fourth-grade level." In a second report, Gates (37) cited additional data of a similar nature for an unspecified number of second grade children of average intelligence.

McElroy (65) examined graphs of median standardized test scores in vocabulary and reading comprehension to determine growth trends for an unspecified number of pupils of above average ability in grades 2 through 7 in an upper-class community. A general growth trend was found in which reading comprehension scores were usually higher than vocabulary scores, but less growth appeared to occur between grades 3 and 4, and between grades 4 and 5, than was the case for other grade levels. By grade six, scores were "up to the expected standard." The author suggested that more challenging materials are needed in grades 3 and 4, especially for the more able students, and made a number of recommendations for strengthening the reading program in the school concerned.

3. Learning Processes and Functions in Reading

Reading Set. Two controlled experimental studies examined the effects of specific learning sets upon reading processes. In an extension of an earlier 1960 study, Muehl (76) conducted a visual discrimination pre-training experiment using differential word sets among 60 kindergarten pupils to determine the stimuli relevant to the recognition of three words. Children ranging in age from 61 to 71 months were randomly assigned to one of three pre-training groups; 1) 15 pupils in a relevant shape-relevant letter, word-discrimination group; 2) 15 pupils in an irrelevant shape-relevant letter, word discrimination group; and 3) 30 pupils in a letter discrimination group used to control the effect of learning relevant letter discrimination upon later word list performance. Nonsense words were used in pre-training, each of the pre-training sets were presented to the word-discrimination groups in four orders, and all subjects received eight matching trials under standardized presentation conditions. In the reading task, word shape and letter differences were adjusted to the appropriate pre-training emphasis. All subjects continued in the reading task to a criterion of 11 of 12 correct anticipations in two successive two-trial blocks, or for a maximum of 20 trials. Analysis of variance revealed no significant performance differences between the two experimental groups of 15 children on either the pre-training or reading tasks, "a finding consistent with the assumption that children discriminate among similar length words of different shape on the basis of specific letter differences." A comparison of the combined word discrimination pre-training groups and the letter discrimination pre-training groups showed "a performance difference favoring the letter discrimination group in pre-training, but no differences on the reading task between the groups." The author concluded: "These findings suggest the hypothesis that although letters presented in words are more difficult to discriminate than the same letters presented singly, this relative difficulty in learning to discriminate letters in words can be overcome by providing visual discrimination with the relevant letters prior to presenting them as parts of words."

Torrance and Harmon (104) investigated the effects of three textbook reading sets upon the test performance of 115 college graduate students. Each student was assigned alphabetically to one of three groups, each of which was given a different reading set: 1) read to remember, 2) read to evaluate, and 3) read to creatively apply. Each group read one week for each of the sets given. Each subject was asked each week to report the degree to which he was able to maintain the desired set, difficulty experienced, time spent in reading, and his semantic reactions to the experience. Each Friday a twenty minute test was given on the day's assignment, the test including cognitive, memory, divergent thinking and evaluation items. The effects of reading sets upon test performance and upon those aspects reported by students, as well as test scores x weeks and weeks x groups, were computed by analysis of variance. Evidence was secured that mature students could assume various reading sets that had differential effects on the kinds of goals achieved, especially for creative thinking and evaluative thinking outcomes. Students were found to maintain their reading sets reportedly only about half of the time, but the sets produced measurable effects on tests covering the reading assignments. Students were more successful in maintaining the memory set and found it easier to maintain, but in only one of three weeks did the memory set result in superiority on the cognitive type of test. This study has provocative implications for the role of reading sets and for alternative ways of measuring reading comprehension at perhaps other than the graduate school level.

Reading and Purpose. Troxel (105) investigated the speed-comprehension relationship and associated factors in the reading of the same expository mathematics materials by 45 eighth grade students, 22 of whom read twenty specially prepared passages averaging 326 words in length to answer one specific question each, and 23 of whom read for the main idea of each. Students were matched on the basis of Iowa Silent Reading Test scores and their reading speed on the experimental passages timed in five-second intervals. While recognizing a number of limiting factors in the experiment, the author suggested a number of conclusions among which were: 1) that purpose influences speed of reading; 2) that a general reading ability test such as the Iowa Silent Reading Test measures "the kinds of things" measured by reading for the main idea in this experiment better than reading for a more specific purpose; 3) that greater accuracy and speed in reading expository mathematics materials tends to be associated with higher general reading ability scores; 4) that speed and accuracy on the timed passages of the Iowa Silent Reading Test is more like that in reading for the main idea in the experimental task; and 5) that faster readers tend to be more accurate in the reading of expository mathematics materials. The findings tend to support those reported recently by Shores on the rate-comprehension relationship in reading science materials for different purposes.

In a college study, Stone (101) examined the effects of training on different types of idea-collection systems upon speed of reading for 62 college students and for a randomly selected control group of 60 students with similar verbal SCAT scores. Directed practice was given primarily in 1) selecting topics which served to organize what was read and 2) in topic selection plus the categorization of details, examples and conclusions. A total of 54 readings of 500-1500 word passages were used during the training period. Students recorded three reading speeds for each session: 1) usual speed, 2) speed for idea-collecting, and 3) pressure speed (reading as fast as possible with comprehension.) Outlines were prepared and evaluated after reading. Improvement in usual reading speed in the experimental group was noted at the five percent level of significance, but speed for idea-collecting improved markedly more, differences being obtained at the one percent level of significance. Correlations among the various speed measures taken at the beginning and end of the training period were likewise cited.

Accuracy of Perception. Harcum and Dyer (45), using tachistoscopic techniques reported in an earlier study of children, measured accuracy of perception to the right and left of a fixation point for 12 adults of unspecified intelligence and reading proficiency. Superior performance was noted in left-hemisphere accuracy regardless of the viewing eye under monocular and binocular conditions, a finding briefly discussed by the authors in relation to the problem of establishing the left-to-right sequence in reading.

Association and Retention in Reading. Otto (81) conducted a controlled experimental study of the ability of a total number of 108 good, average and poor readers in grades 2, 4, and 6 to learn and recall a list of paired associates by three modes of presentation. All subjects were in the 95-108 IQ range and within six months of average chronological age per grade. Good readers were selected from the upper three stanines, poor readers from the lower three stanines, and average readers from the middle stanine of scores upon the Metropolitan Reading Test. Simple geometric forms were presented in a Stone memory drum for a four-second interval, and subjects paired low association value consonant-vowel-consonant trigrams with the appropriate forms to a criterion of one correct anticipation of the entire list. The three modes of presentation of the trigrams by the experimenter were auditory (saying the trigram), visual auditory (seeing and saying the trigram), and visual-auditory-kinesthetic (seeing and saying the trigram, plus having the pupils trace the trigram on a card.) The subjects re-learned the list in scrambled order after 24 hours, and the total number of trials for the original and the re-learning tasks was taken as the total score. An analysis of variance was employed to study possible interaction effects for the variables studied, significance being accepted at the one percent level of confidence. The investigator found 1) that the groups of good, average and poor readers, in that order, each took significantly more trials to perform the experimental task; 2) that sixth, fourth and second graders, in that order, similarly and significantly took more trials, but significant interactions were found to involve reading level and mode of reinforcement; 3) that mode of reinforcement interacted significantly with grade level; 4) that there was not a significant interaction between mode of reinforcement and reading level "contrary to the assumption in instructional practice that additional cues will be of value in a remedial teaching-learning situation."; and 5) that there were no significant differences between good, average and poor readers in the number of trials required to re-learn the list. The finding that the effect of mode of reinforcement appears to be different according to grade level but not according to the reading level of the pupils should be of particular interest to those concerned with remedial reading.

Rothkoff (85) investigated the retention effects of order of presentation of topically related but loosely connected sentences comprising three passages, as well as the location of information within sentences, under controlled time conditions for 144 female switchboard operators. The passages reportedly represented three levels of difficulty. A counterbalanced design was used in exposing the three passages to each of 12 groups of 12 subjects. Retention was measured after one and 15-minute intervals by sentences in which one word was omitted. Among the findings reported were the following: 1) the three passages produced three different performance levels; 2) retention loss during the 15-minute interval was small; 3) the serial order of sentence presentation had little or no effect on retention; and 4) a serial order effect significant by Chi-square analysis at the one and two percent levels of confidence, respectively, was associated with the location of the test sentence in the testing sequence for two of the three passages.

Cohan (20) reported that a two and one-half year old girl of average motor and language development and behavior successfully learned to discriminate between words printed on cards and remember these over a six weeks' period, thus adding another case report to the many on early reading performance.

Feedback of Oral Reading. Two studies illustrate the great dependency of the reader upon immediately hearing his own voice for adequate oral reading performance. Using a delayed sidetone technique in which the interval between speaking and hearing one's voice may be controlled through the use of electronic controls and headphones, Schwartz (89) asked 10 male college students to read by whispering each of of 11 lists of 15 five-syllable phrases equated for mean intensity and duration values under different intervals of delayed feedback. An examination of the mean length of time taken to whisper the lists for delay periods, adjusted by .03 second intervals and ranging from 0-.21 seconds, showed that subjects required more time as the amount of delay was increased, a finding determined by analysis of variance to be significant at the one percent level of confidence. In a second investigation, Garwood and Boyer (35) experimentally studied the length of time 80 college undergraduates took to read orally three types of emotionally-toned sentences under conditions of delayed and non-delayed sidetone feedback. Eighteen sentences were used-six of which were judged to be pleasant, six unpleasant, and six neutral - with approximately the same number of syllables per sentence. Each subject read the sentences first under conditions of non-delay and then under conditions of a side tone feedback delay of .20 seconds. Analysis of variance for the combined conditions showed that delayed feedback produced significantly longer reading times at the one percent level of confidence; similarly, that subjects took a significantly longer time to read the pleasant and unpleasant than neutral sentence groups. The authors concluded that speech content is a significant source of variability with respect to oral reading time.

Oral vs. Silent Reading Effects. Collins (22) investigated the effects of oral vs. silent reading for 60 college freshmen upon the comprehension of seven 300-word stories written at seven levels of difficulty according to the Flesch reading ease formula. The 30 students who read silently were matched for age, sex, SCAT scores, and the Stanford Reading Comprehension Test with the 30 who read orally. The difference in mean comprehension scores on each selection varied but consistently favored the oral reading group, while the difference for the total comprehension score was significant beyond the one percent level of confidence in favor of the oral reading approach.

Aural vs. Text Presentation Effects. Newton (77) conducted an experiment to determine the effects of personal preference in relation to reading vs. presentation of the same technical material to groups of airmen in basic training. After a pretest, subjects were asked their preference for a reading presentation, an oral presentation, or no preference. The airmen were then put in a group of their choice, and randomly assigned to subgroups which were given the same material, one written and one orally. The materials used were seven 200-400 word paragraphs with a two and one-half minute reading time limit. Post-test results were analyzed by analysis of variance for choice, mode of presentation, and ability level (above or below mean AFQT score). No significant differences were found in test scores for the three preference groups. While evidence was presented that possible significance at the five percent level of confidence obtained for the brighter group, for the reading mode of presentation, and for the interaction level x mode and choice x mode, the author concluded that whether or not such subjects get their favored mode appears to be of relatively little importance.

Listening and Reading. Does practice in listening in an upper elementary grade affect reading achievement? Lubershane (60) reported the relationship of a program to improve listening to reading achievement for an experimental group of 35 and a control group of 37 fifth grade pupils. Groups were matched for chronological and mental age, subject-matter achievement, and reading achievement upon the Metropolitan Reading Test. Upon re-testing, the results favored the experimental group in reading, a finding supporting the position taken by Cleland and Touissaint (19). The author concluded: "Although there is no definite statistical proof of the values of the exercises in improving reading ability, the generally greater growth in reading ability by the experimental growth suggests strongly that the auditory exercises have a positive effect on reading growth."

Reading and Programmed Learning. Calvin and Hanley (17) reported briefly on the effects of programmed instructional material upon reading rate and level of comprehension for an experimental group of 215 students and a control group of 156 students enrolled in high school mathematics courses. End of the year results on the Cooperative Test of Reading Comprehension showed that the experimental group was significantly better than the control group in reading rate at the five percent level of confidence, but there were no significant differences in vocabulary or level of reading comprehension.

4. Research in Readability

Reading Age and Readability. McLeod (67) used the known reading levels of children in grades 2 to 4 to estimate the readability of books 1, 2, 3 and 5 in the Blue Book series of the *Wide Range Readers* used in England. "The reading age at which 50% of the children could read a sample (with not more than two mistakes per 30-word passage) was defined as the difficulty level of the book." In the first experiment, a word recognition criterion was used for 255 children; in the second, comprehension of a factual, literal question was the criterion for 375 children. Comparative data were reported for readability according to the Dolch and Spache formulas, the author's estimate, and for word recognition and comprehension reading ages of the children. The author reported that the readability levels of the books according to the Spache and Dolch formulas bore no consistent relationship to the experimental findings nor to each other.

Effect of Shortening Sentences. Coleman (21) examined the question of whether or not shortening sentences improves comprehensibility by preparing three versions of the same material with average sentence lengths of 39, 23 and 15 words and using the cloze test procedure for measuring the readability of the nine resulting different selections for 90 college undergraduate students. By dividing the subjects into five groups and constructing five different cloze tests for each passage, each with a different selection of words deleted, a score was obtained for every word. The sum of scores for all words in the passage was the score for the passage as a whole. Subjects were given 50 seconds to read each passage and were tested only on the portions completed. Each subject read one passage in each style, but different subjects read different style-passage combinations. It was found by "t" tests that the shortest version was significantly more readable than the longer ones but the advantage was small, being only seven percent more comprehensible. The author suggested that shortening a clause may be more important than emphasizing the boundaries with a period and a capital.

Applications of Readability Formulas. The Dale-Chall readability formula was used in several studies. Groff (42) analyzed the readability of samples taken six months apart of two issues of 14 children's magazines. Wide differences in the readability of various magazines were found, ranging from grade levels of 2.1 to 6.8. Differences between two issues of the same magazine ranged from .1 to 4.6 grade levels for individual articles, with average differences of from 0.0 to 1.2 grade levels. The author reported that editors said few or no controls were used to make sure the vocabulary was not too difficult, but relied chiefly on the experience of the editors. The magazines analyzed are listed, most of which have readability levels appropriate for intermediate grade use.

Belden (5) studied the relationship of the readability of five high school biology textbooks measured by the Dale-Chall formula and reading scores of 357 tenth grade biology students on the Nelson-Denny Reading Test. Accepting a suggested criterion that the reading achievement levels of students using a book should be one grade above the difficulty level of the book, the author found that the five books met this criterion for the following percentages of students per book: 37.0, 39.5, 42.3, 49.6, and 58.5 percent. He concluded that only one of the five textbooks was appropriate in terms of its difficulty in relation to student reading achievement for over one-half of the student. Nyman, Kearl and Powers (79) reported that the predictive power and error of the Dale-Chall readability formula was seriously affected when the basic word list used in the formula was shortened from 3000 words to 920 words.

The Physiology of Reading

Specific Dyslexia. Since the term "dyslexia" has a strong neurophysiological connotation and is the subject of a recent volume entitled Reading Disability, the latter may be considered here as a contribution to the present state of knowledge of the physiology of reading.

Money (75) edited and provided a summarizing statement concerning 12 papers presented at a 1961 conference at Johns Hopkins University to explore the present status and research needs in "specific dyslexia as it affects the learning of reading." The scope of authorship, topics and approaches represented is broad, the subject being considered as "a pedagogical problem, an epidemiological problem, as a clinical problem in medical psychology and neurology, and as a problem in experimental research." An introductory statement by Leon Eisenberg is particularly valuable in clarifying the type of reading disability to which the conference members applied this term and in candidly indicating the present state of knowledge of specific dyslexia. "It is a matter of considerable importance to be aware of the different senses in which the concept of reading disability is employed. .... To begin with, there are those who cannot read because of a sensory handicap, as in the case of extremely poor vision or hearing. Next there are those whose inability to read is part of a general pattern of inability to learn, namely, the mentally deficient. The third group is made up of those children who have had deficient or improper instruction by virtue of insufficient or poor schooling. A fourth group can be said to be made up of those children who, despite an educational opportunity, are so poorly motivated to learn to read because of circumstances at home that they continue to lag behind their grade norms. It is my view that for none of these groups would the term, 'specific dyslexia,' be appropriate. For I would apply this term to a situation in which a child is unable to learn to read with proper facility despite normal intelligence, intact senses, proper instruction, and normal motivation. It is evident that in this construction the term 'specific' really implies an idiopathic condition, that is, one whose cause is unknown." (pp. 4-5) The task the conference set for itself was to explore the latter condition and in so doing provide a useful perspective to the study and understanding of a difficult and complex problem, particularly with reference to current theories of reading disability.

Deanol, A Psychic Energizer. Two studies reported the effects of the use of Deanol, a psychic energizer, in relation to reading improvement programs. Staiger (96) explored its use in 60 pairs of retarded readers matched for age, reading scores on the Gates Reading Survey, and intelligence, drawn from elementary, high school and college populations. Deanol was administered in two 75 milligram doses per day to the experimental group for eight weeks, while placebos were given the control group. Perceptual speed was measured by the Clerical Speed and Accuracy subtest of the DAT. Analysis of variance revealed no significant differences between the Deanol and placebo students except in clerical speed and accuracy in which the Deanol subjects improved in all but one comparison beyond chance expectancy. The author noted that reading was not affected to any great extent by the use of Deanol, although there was some increase in the F values of reading sub-test scores favoring the experimental group. Deanol was similarly administered in a study reported by Shabaglian (90) to experimental and control groups of 30 mentally retarded boys with IQ's of 50-69 for an eight-week period in connection with a two-hour daily reading program. Reading progress was measured on the Gates Primary Reading Tests and classroom quizzes. Since no significant differences were found between the means of the quizzes and the silent reading tests used, the author concluded that Deanol did not improve learning to read in this study.

Visual Dysfunction and Reading. Spache and Tillman (95) compared differences in the means and the percent of failure on Ortho-Rater visual test scores of 114 retarded readers with 101 normal readers of comparable mental ages in the elementary school. While it was found that the two classes of readers were more alike than different, there was evidence 1) that retarded readers had lower left-eyed acuity at the near point, significant at the five percent level of confidence, and 2) that there were significant differences at the one percent level of confidence in the proportion of retarded readers failing the near-point acuity test and in the number of this group having differences in the near-point acuity of the two eyes. These findings were interpreted by the author as related to interference with comfortable, accurate binocular vision and to fusional difficulty frequently associated with reading difficulty.

Ewalt (30) administered tests of functional visual performance to 50 retarded readers of various ages and intellectual levels, most of whom had 20/30 vision or better and all of whom had passed the Snellen test for distance vision. It was found that the majority had poor visual functioning and inadequate eye movements, and that almost a third had a bilateral constriction of the form field of 20 degrees or more.

Gruber (44) reported little relationship between the findings of orthoptic analysis and opthalmograph records of binocular coordination for 50 randomly selected patients, aged 10-68 years.

"Reading Epilepsy." Lasater (57) contributed four case study reports of adults who experience epileptic seizures in the jaw only while reading and have done so since the onset of adolescence. According to the author, 20 other such cases have been reported previously.

The Teaching of Reading

1. Status of Reading Instruction

Elementary. An earlier study not previously reported in this summary is pertinent. The British Ministry of Education conducted surveys in 1952 and 1956 of the reading attainment levels of eleven and fifteen year-old children in England to determine if there had been any change from the standard reported in a previous 1948 survey of England and Wales. The comparative results are published in a pamphlet entitled Standards of Reading, 1948-56 (98). The 1952 and 1956 surveys were marked by increasingly more refined methods of stratified sampling described in some detail in an appendix. The same reading test devised by Watts and Vernon for the 1948 survey was used in the later surveys, consisting of a 30-question, multiple-choice power test of reading comprehension with a ten minute time limit. Case study data were also gathered on students who did exceptionally well or exceptionally poorly on the test. While the study should be consulted for detailed findings, the general results showed 1) that "In the interval from 1948 to 1956 there was on the whole an improvement in the ability of school children to read with understanding."; 2) that there was a nine month advancement in the mean reading level of 11 year-old children and a five month advancement for 15 year-olds during this period; and 3) that the proportion of "really good readers" had nearly doubled from nine to 17 percent, while that of semi-illiterate and illiterate children had shrunk from five to one percent. The study provides evidence from another country that reading attainment in today's schools is generally higher than that in previous years, as has been reported for American schools by Gates and by Worcester and Kline.

The status of an enriched basal reading program was described by Brown (15) who cited test results for approximately 1500 pupils in each of grades 3, 5 and 8 between 1957 and 1960 upon standardized reading tests to illustrate the effectiveness of a basal reading program supplemented by a wide range of independent reading in a "stable conservative community." It was found that mean scores on the California Reading Test in grade 3 and on the STEP reading test in grades 5 and 7 were well above national norms.

Secondary. McGinnis (66) compiled the results of questionnaires completed by 570 out of 1000 teachers contacted, and by 1029 college freshmen, concerning various aspects of secondary reading instruction. The results of the teacher questionnaire, too detailed for mention here, support the findings by Strang and others concerning the need felt by secondary teachers for the improvement of secondary reading and the inadequacy of their preparation for this task. The majority of the college students reported that they had not been helped to improve their reading in high school, that teachers required all to do the same work and read the same text, and that they had not been taught how to concentrate nor to read a chapter effectively.

Grissom (40) conducted a follow-up study of the characteristics of 15 secondary reading improvement programs, representative of an earlier survey of 107 such programs in Illinois schools, using questionnaires, observations and interviews to gather salient data. The presence of three types of programs identified in the earlier survey was confirmed, while programs were found to be characterized by cooperative staff efforts in selecting students, in providing an appropriate atmosphere, and in adjusting to individual differences. Larger schools were found to have more flexible program patterns, frequently involving an all-school reading clinic, than small schools. Geake (38) conducted a questionnaire survey of existing and proposed special reading programs in Michigan high schools. Returns were analyzed from 90% of a systematic sample of 128 schools by taking every fifth school from an alphabetical listing according to four classes of school size. The results showed 1) that by the end of 1961, 50% of the schools would have some special instructional arrangement for teaching reading; 2) that proportionate figures for 1955 and 1961 revealed an increasing percent of reading programs for schools under 500, the reverse being true for the larger schools which initially reported a higher percent of programs in 1955; 3) that a slightly higher percent of the larger schools had programs in the planning stages than did small schools; 4) that special classes with a special teacher was the most typical administrative arrangement; and 5) that most respondents were favorable toward such programs, rapid reading classes and a combination of programs rating the highest in satisfaction in that order.

The status of a leisure reading program in Canadian secondary schools appeared to be the particular objective of a study by Fisk (34). A questionnaire survey was made of 11,000 Canadian students and 2000 of their teachers in junior and senior high schools concerning reading habits, interests and attitudes, and the adequacy of school facilities and teacher preparation for leisure reading. Among other things, the findings indicated 1) that teachers were more conservative than students in estimating the number of books read by students and the degree of pleasure they derived from reading; 2) that there was evidence of a great shortage of suitable books typically available for leisure reading; 3) that there was little evidence of an excellent variety in the leisure reading diets of students; 4) that the recognition of leisure reading as a bona fide activity in English programs was questionable; and 5) that teachers felt unprepared to conduct an effective leisure reading program. On the basis of these and other specific findings, the investigator concluded that the leisure reading program in the schools surveyed definitely needed to be improved in material ways and in the preparation of teachers for such a program.

College and Adult. Colvin (23) described the characteristics of an "ideal" college reading program derived from answers to a questionnaire returned by representatives of 42 colleges and universities in Pennsylvania. The recommendations with respect to the need for such a program and its essential content and approach are reflected in many existing college reading programs, but the recommendation that all freshmen be required to take the course at one of four levels would be challenged by those who favor voluntary programs, possibly including Feinberg, Long and Rosencheck(31).

At the adult level, Acker (1) cited further data gathered from the analysis of a survey of 177 adult reading programs in military, business and government agencies reported earlier. He found 1) that the objectives of increasing vocabulary and reading comprehension were prominent; 2) that 25 hours was the typical reading program length with most programs emphasizing mechanics of reading; 3) that two-thirds of the programs were voluntary while about half limited attendance to executives, supervisors and selected employees; 4) that 60% of the agencies used group instructional methods, 30% combining this with individual instruction; 5) that while all used standardized tests, 64% developed non-standardized tests to meet local needs; 6) that typically there was only one instructor, and of these 43% had one year or less of training in the field of reading; and 7) that reading proficiency tended to improve reading speed, vocabulary and comprehension in that order. The author concluded that such programs were effective and needed.

2. Developmental Reading Instruction

Grouping. Groff (43) analyzed changes in basal reading instruction groups in grades 2, 3 and 4 as reported by student teachers in 66 classrooms for 2179 pupils. Changes were reported weekly from September to January during a school year. Specific findings included evidence 1) that 365 changes were made during this period, 55% involving boys and 45% involving girls; 2) that 10% of the teachers made no changes, 39% made 1-4 changes, 32% made 5-9 changes, 12% made 10-14 changes, and 6% made 15-17 changes; 3) that changes were made most frequently in grades 2, 4 and 3, in that order, averaging about one-third of the total number of changes made; 4) that there was a larger percentage of change in the first weeks than later; and 5) that the number of pupils in the low reading groups was smaller than the middle or high reading groups. It should be noted that the total number of changes was relatively small in proportion to the total number of pupils in these classrooms, a finding which may be interpreted in different ways regarding the efficiency and flexibility of the grouping practices reported.

Evaluation of "Easy-to-Read" Books. Russell (86) used subjective analysis of content and motivation, objective analyses of vocabulary and readability, and comments by first and second graders and their teachers in evaluating 10 "easy-to-read" trade books for children. The books represented well-known authors and various categories of content. Among the findings reported were these: 1) that readability as measured by the Spache Readability Formula ranged from 1.7 to 2.9 in grade level placement; 2) that the books corresponded in difficulty estimated upon several bases to the first readers and level-one second readers from two well-known basal reading series; 3) that any one trade book provided about the same amount of reading material as a second basal pre-primer; 4) that the books were generally attractive in format and were judged suitable for young readers ranging from superior readers in grade 1 to below average readers in grade 2; and 5) that the books were generally well received by teachers and pupils. It is perhaps significant that in a month's classroom trial these books were read mostly by pupils in grades 2 and 3, while first grade pupils did little reading of them.

Comparative Instructional Methods. Blakely and Shadle (8) reported a comparison of two approaches to reading readiness in kindergarten, the use of readiness books of a basal reading program vs. an experimental activity program of experiences, for two groups of 28 children during one school year. The groups contained an equal number of boys and girls, were taught by the same teacher, and were found to be not significantly different at the beginning of the study on the Metropolitan Readiness Test, a maturity checklist and an informal reading readiness appraisal. It was found 1) that all mean gains favored the experimental activity group; 2) that significant differences were noted at the one and five percent levels of confidence, respectively, for the experimental group upon the maturity check list and the informal appraisal but not on the Metropolitan Readiness Test; and 3) that boys, but not girls, made similarly significant gains in the appraisals made. The authors, while recognizing several limitations of the study, suggested that the activity approach is to be preferred especially in view of the gains made by the boys.

Studies of individualized vs. conventional methods were reported by Aronow and by Jones and Van Why. Aronow (3), after identifying 351 sixth grade pupils in New York City who had received individualized reading instruction in grades 4 and 5, compared their reading performance with 6729 sixth grade pupils who had not received such a program of instruction. The two groups were shown to have been comparable in third grade according to the Otis Alpha (a test of non-verbal mental ability) and the Metropolitan Reading Test. The pupils in the individualized reading program scored significantly higher at the one percent level of confidence in their mean score on the latter test than did pupils in the regular program. The implications of these findings were discussed at some length. Jones and Van Why (52) compared the mean vocabulary and reading comprehension gains on the Iowa Tests of Basic Skills for 52 experimental and 52 control pupils matched for reading performance and intelligence and given differentiated reading instruction for five months in the fourth grade. The experimental group used the SRA Reading Laboratory materials. Although both groups made significant gains, there were no significant post test differences between the groups. Within-group differences significant at the five percent level of confidence were noted in the experimental population, high achievers making the greatest gains in vocabulary and low achievers making the greatest gains in reading comprehension. While many factors may account for the differences in the findings of these two studies, an important one may be that of the relative length of the programs and its associated variables.

The use of instructional film-strips continues to receive attention. Kelly (54) reported the effects of using film-strip presentation of basal text material and lessons on phonetic and structural analysis from the teacher's manual for two experimental and one control groups involving a total of 138 children in the second semester of the first grade. Grouping was used in only one of the experimental groups, and reading progress was measured on alternate forms of the Gates Primary Reading Tests. The results of the experiment indicated 1) that the film-strip groups were significantly superior to the control group on mean scores for word recognition, sentence reading and total reading scores but not on paragraph recognition; and 2) that all test differences for boys but not for girls were significant. Data on the initial comparability of the groups and on the level of confidence accepted in the analysis were not specified.

The question of integral vs. supplementary isolated phonics instruction was investigated by Ibeling (49), who compared reading and spelling test gains among randomly assigned experimental classes in grades 2, 4 and 6 which received supplemental phonics workbook instruction with classes not receiving such supplemental instruction. The 600 pupils involved in the study, about half of whom were in the experimental group, represented various socio-economic levels, the majority being from the upper or middle class. The time spent on reading and spelling was reported to be about the same for both groups. The phonics workbook Building Reading Skills was used in grade 2, and Phonics We Use in grades 4 and 6. Among the results of the study it was shown 1) that no consistent advantage obtained for either experimental or control groups on post-tests of reading and spelling achievement; and 2) that of 15 analyses made, only two showed mean differences favoring the experimental group at the one percent level of confidence: visual analysis on the Silent Reading Diagnostic Tests and spelling achievement in grade 2. The results of the study properly raise the question of the value of purely supplemental phonics workbook instruction, indiscriminately applied to all pupils whether they need it or not at the grade levels concerned, as do those of a remedial reading improvement program reported by Love (59).

The role of reading workbooks was likewise explored. Docter (26) examined the comparative effectiveness of the use of reading workbooks vs. non-workbook activities in experimental and control groups in 36 classrooms from 18 randomly selected elementary schools in Los Angeles, and likewise reported questionnaire opinions on the same topic by 1200 teachers and 89 school administrators. While specific data are not cited for the experimental study, the author reported significant post-test differences on SRA reading comprehension and vocabulary tests favoring the non-workbook group in grade 1, favoring the workbook group in grades 2 and 3 and in comprehension only in grade 4. Post-test differences were not significant in grades 5 and 6. Results of the questionnaire findings, further evaluated by a city-wide committee of 60 teachers and administrators, tended to support the use of reading workbooks except for reactions by first-grade teachers. The author concluded that the use of reading workbooks was more of a boon than busywork, being a challenge to children reading above grade level but frustrating to children reading below grade level.

An extensive study of the effects of speech training upon reading was conducted by Sommers et. al. (93) who investigated the effects of speech improvement classes of varying lengths upon the reading improvement of large groups of children in grades 1 and 2. Experimental and control groups were used to measure the results of speech training programs of the following lengths: 1) 16 weeks in grade 1; 2) nine months in grade 1; and 3) eight weeks in grade 2 following the nine months in grade 1. The weekly speech improvement period was 30 minutes. Pupils were given the Primary Reading Profiles Test and measured upon the combined reading factors of reading aptitude, word recognition, auditory association, word attack, and reading comprehension. Results of analysis of variance showed that second grade experimental subjects made significantly greater scores at the one percent level of confidence on the combined reading factors, but not on reading comprehension, than did their controls. There were no significant differences in the reading performance of experimental and control groups at the end of first grade.

3. Reading Improvement and Remediation

Many studies reported under the topics Psychology of Reading and Physiology of Reading are relevant here.

Special Reading Programs. Brueckman (16) reported the reading status of 96 juniors and seniors in an advanced reading program according to scholastic aptitude to show the relatively high and consistent levels attained in reading, and in reading rate gain, by honors students and those of average scholastic ability. Mean reading comprehension scores of 12.2 were reported for juniors and of 13.5 for seniors. North (78) cited mean and median gains on the Diagnostic Reading Test for 368 high school seniors as a result of 27 periods of work in a reading laboratory, arranged in two sessions spaced six weeks apart. Mean pre-and post-test scores in percentile rank on national norms were: 1) rate, 90-96; 2) comprehension, 63-67; 3) vocabulary, 51-61; and 4) total reading score, 58-65.

At the college and adult levels, Magoon and Magoon (69) described the academic abilities of students in a college reading laboratory and cited the progress of college freshmen as a function of the number of practice sessions attended. Students participating in the reading laboratory were found to be distributed rather evenly among the quartiles on the ACE according to local norms. Some evidence based upon a limited and variable sample was offered that gain in mean reading speed and possibly mean change in percent of comprehension was roughly proportional to the number of practice sessions attended. Fields (33) described the nature of a mandatory six weeks reading program taken by freshmen as a part of the English curriculum in four high schools, illustrated gains made in the program, and outlined plans for the expansion of the program. Siegel (91) described the population characteristics and results of an adult reading program conducted with the collaboration of a local college staff over a five year period in a library setting. Different phases in the organization and conduct of the program are described and a number of suggestions are offered to those interested in adult reading programs.

The question of drop-outs in college reading programs was likewise investigated. Hinton (48) reported on the drop-out rate and scholastic averages of 71 college students in a non-credit college reading improvement program and 24 honors section students. The drop-out rate was found not to be substantially different. The records of 47 reading improvement students and 18 honors students were then examined in relation to their scholastic averages in previous, current and following semesters. It was found that there were differential effects between the groups in gains in grade point means and medians, the reading laboratory group improving from the previous to the current semester but the honors group decreasing from the previous to the following semester, a degree significant for both groups at the five percent level of confidence. Wood (111) examined the extent to which personality variables and selected instructional variables might affect attrition in a representative sample of 701 students in a non-credit college reading improvement program of six to seven week's duration. The first hypothesis that attrition is associated with measurable student personality and motivational characteristics was explored among all registering students by 1) the 16 PF test; 2) SA-S scales developed earlier from questionnaire items differentiating slow and fast improvers in the reading improvement classes; and 3) inferences about motivation made from advisor referrals and academic need. The second hypothesis that instructional variables are associated with attrition was explored by 1) efficiency on daily practice exercises; 2) student ratings of instructors; 3) questionnaire regarding student friends in same or other reading improvement class; 4) time of the reading class; 5) experience of the instructor; and 6) questionnaire response by "drop-outs" on reasons for discontinuing the class. The first hypothesis was not confirmed, but the second received moderate confirmation. The author concluded that the proportion of attrition in such a course is a legitimate criterion for evaluating the general effectiveness and quality of the program.

Comparative Studies of Reading Improvement. Schrock and Grossman (88) used visual training exercises accompanied by "encouraging remarks" about reading as a motivational device to study its effects upon gains on the Gray Oral Reading Test for 20 seventh grade students reading two or more years below grade level vs. a similar control group of 10 students. Both groups were from the 90-110 IQ range on the California Test of Mental Maturity. A blind analysis of results was conducted and gains appraised by testing for the significance of the difference of means and by Chi-square analysis. Both appraisals yielded highly significant differences favoring the experimental group at the one-half percent (.005) level of confidence. While the authors noted a high incidence of various visual dysfunctions according to the

Keystone Visual Survey Tests among their population, they concluded that motivational factors were paramount in explaining the superior progress of the experimental subjects. Warren (108) compared the results of a seven-week junior high school reading program using reading instrumentation for 25 eighth grade students with a comparable program and number of students not using instrumentation. The classes were equated for IQ and scores upon the Iowa Silent Reading Test, and eye movement photographs were taken to measure "functional reading skills." The teachers of each group were judged to be of equivalent ability and conducted parallel classroom presentations. Among the results of the study it was found 1) that the mean gains on the Iowa Silent Reading Test favored the experimental group; 2) that the rate subtest scores on the latter test were significantly better for the experimental group at the one percent level of confidence; and 3) that eye movement photography showed differences favoring the experimental group in number of fixations and regressions at the five percent level of confidence, and in scores for rate-comprehension and for relative efficiency at the one percent level of confidence. Love (59) compared the progress of two groups of 15 children averaging 12 years of age in a six week remedial reading program in which one group received phonics instruction as an integral part of an individualized reading program while the other received separate drill in a phonics workbook which employed a whole word approach. The groups were equated for intelligence on the WISC and for reading achievement on the Gates Reading Survey. The results showed mean reading gains for both groups but a negligible difference in amount of gain, a finding comparable to that of Ibeling (49) in a developmental study.

At the college level, Feinberg, Long and Rosenheck (31) compared the mean reading achievement of 64 college freshmen in a mandatory reading and study skills program in 1960 with that of a matched group not enrolled in a mandatory course in 1959. Groups were matched for sex, high school type, high school average, and verbal and mathematics SAT scores. No significant change was found on the Cooperative Reading Comprehension Test as a result of the reading program except for a significant loss in vocabulary. The negative results of the program were attributed to resistance toward the mandatory nature of the course, thus offering an interesting contrast to the views of college personnel on the ideal college reading program reported by Colvin (23). Bloomer (12) studied the relationship between reading improvement on the survey portion of the Diagnostic Reading Test and academic achievement for 40 experimental and 39 control subjects in a twelve-week college reading improvement program. Among the findings it was noted 1) that by "t" test analysis of the difference of mean scores the experimental group made significantly greater gains in reading speed and comprehension, but that the control group similarly exceeded the experimental group in vocabulary gains; and 2) that gains in reading skills were not found to be related to academic achievement nor to gains in academic achievement. When grade point averages were corrected for predicted grade point average, however, the experimental group was found to have gained significantly more. The author concluded that other variables than reading affect high academic performance. In a second study, Bloomer (10) investigated the effects of the use of the cloze procedure as a remedial reading exercise in an experimental group of 44 college students in comparison with a control group of the same number enrolled in a college non-credit reading improvement course. The experimental subjects were given a series of 500-word passages from basal reading, science and social studies programs for elementary school use in that order, in which every tenth word of a passage was omitted. Reading levels and an error criterion were set for students to meet to progress to more difficult material. Comparisons of pre- and post test means on the Diagnostic Reading Survey Test showed that the cloze procedure exercises were associated with increases in the reading comprehension and total reading scores in the experimental group to a degree significant at the five percent level of confidence, and similarly in the average college grade predicted. The author felt that the cloze exercises were more motivating than the conventional materials used by the control group, for the latter group, while gaining significantly in vocabulary, showed an equally significant loss in comprehension in this study.

Remediation of Severe Reading Disability. The relative merits of group vs. individual instruction for disabled readers were the interest of two British investigations. Johnson and Platts (51) compared the reading achievement scores obtained one year after remedial teaching with prior data for 74 children given individual remedial reading instruction and 210 receiving group instruction over a five year period at an English Child Guidance Centre. Correlations among several pertinent variables were also secured. Among the findings of interest were the following: 1)that there was little difference between individual and small group remedial instruction either during such instruction or on delayed testing in mean reading quotients, suggesting that the latter is to be preferred because of its economy in terms of teaching manpower; 2) that those who made the most progress during the period of remedial teaching tend, with some exceptions, to have the highest reading ages on the follow-up testing; and 3) that there was a significant relationship between progress made and length of time elapsed since remedial teaching ceased, thus suggesting that maturation may contribute to the findings during the follow-up testing. Keating (53) reported the effects of group and individualized instruction for boys seriously retarded in reading between the ages of 11-4 and 16-2 in an English Hospital School for epileptic children. In Phase I of the study, a group of 20 was taught but most of the group made less than "normal" progress according to a gradient analysis in which rate of improvement was obtained by dividing the increase in reading age by the instructional time. In Phases II and III, children were given 15-20 minutes per day of undivided teacher attention, described as "reading at mother's knee," plus additional time for silent reading or for doing exercises. The latter phases emphasized individualized reading, the third phase being conducted in four groups of five boys each for daily work. Gradient analysis revealed that progress under the more individualized conditions was much greater for more boys than the large group procedure, a result somewhat at variance with that found by Johnson and Platts.(51)

A multiple disability problem was reported by Ancevich and Payne (2) who described the nature and results of remedial treatment of a left-handed, nine year old boy of normal intelligence with multiple language and behavior difficulties, including retardation in reading and handwriting. Since the boy used a writing style that prevented his viewing of words as he wrote, his writing orientation was adjusted to permit more ready access to visual stimuli in writing production on the assumption that this would facilitate associative processes in word recognition and spelling. In learning to spell three lists of 15 new words presented tachistoscopically, it was found that the new writing orientation produced significantly better spelling at the one percent level of confidence than did the subject's usual mode of writing or one which restricted the view of the whole word artificially. Pre- and post-test results for a five-month period on the Durrell Analysis of Reading Difficulty showed marked gains in word recognition and half-year gains in word analysis and spelling, but no gains in other aspects of reading measured.

How shall progress in reading clinic programs be reported? Bliesmer (9) reported data on the effective reading progress of groups of 40 and 41 children from grades 4 - 8 who were enrolled in remedial reading clinic programs over a nine month period. The data were analyzed by grade and IQ groupings with respect to average remedial gain, average yearly gains before remedial work, ratios of average remedial to average yearly gains, and average potential-achievement gap decreases. The relative merits of each method of reporting remedial progress were discussed at some length.

4. Appraisal of Reading

Many of the studies included under Psychological Correlates of Reading are likewise relevant to this topic.

Formal vs. Informal Methods. Kermoian (55) investigated the relationship between informal judgments by 13 teachers of reading readiness and scores on the Metropolitan Readiness Test for 276 first grade children. Teachers used the same five point scale as found on the test during the first two weeks of school. A correlation of .73 was secured, significant at the one percent level of confidence, 85% of all teacher estimates falling within one rank above or below the test ranking. When teacher rankings deviated from the test scores, they usually overestimated rather than underestimated reading readiness. At the sixth grade level, McCracken (63) compared the vocabulary and reading comprehension scores of 56 sixth grade pupils on the Iowa Tests of Basic Skills with ratings on an informal reading inventory. Three informal rating levels were used: 1) immediate instructional level, 2) maximum instructional level, and 3) word recognition instructional level, for which correlations were obtained between the first two and the Iowa Reading Comprehension of .78 and .74, respectively, and between the latter and the Iowa Vocabulary of .78. While these correlations were fairly substantial, the author indicated that reading placement on the basis of standardized tests would place 63 percent of the pupils at their frustration level and 93 percent at a level too hard for them.

Diagnosis of Oral Reading. McCracken (62) used four graded paragraphs from the diagnostic tests for grades 1 to 3 of the Sheldon Basic Readers to analyze oral reading performance of 36 second grade pupils judged good, average or poor readers. It was found that the number but not the type of error was discriminating, and that comprehension appeared to vary without regard to the number of oral reading errors. Observations were offered by the author concerning the extent to which oral reading and the comprehension associated with it should be considered general or specific skills. Lawson and Airla (58) reported briefly on the correlation of two reading tests, the Gray Oral Reading Paragraphs and the reading section of the Wide Range Achievement Test, for 30 mentally retarded adults. A significant correlation of .94 was obtained which led the authors to conclude that there was little difference between the tests, but that the latter test may be preferable because it may be given in less time.

Timed vs. Untimed Vocabulary Tests. Boag and Neild (13) supplied additional data on the power vs. speed controversy in testing by comparing the scores on the vocabulary section of the Diagnostic Reading Test under timed and untimed conditions for 722 students in grade 9-12 English classes. It was found 1) that the speed and power scores were not interchangeable, 2) that changes in the relative scores were greatest in average and non-academic groups, but not sufficiently great enough to benefit students in the latter group; and 3) that the slow but accurate average student benefited from the untimed test.

Concluding Statement

Most of the empirical investigations cited have appeared as journal articles in several general classes of publications which reflect the broad scope of the Summary. Analysis of these publications sources by broad categories reveals this relative ordering: 1) pedagogical journals designed primarily for teachers, 2) psychological journals, 3) educational research journals, 4) sociological journals, including the field of journalism, and 5) medical and physiological journals. While the order of this ranking has not changed from that of the previous Summary, a trend toward a higher proportion of studies appearing in pedagogical journals is evident. A previously mentioned trend toward improved experimental research design is likewise noticeable in a number of studies reported, as investigators identify and study problems in reading in improved ways.

BIBLIOGRAPHY

1. Acker, Ralph S. "Reading Improvement in Government and Business," Education, LXXXII (March 1962), 428-31.
   Analysis of specific data on the operation of 177 adult reading improvement programs in military, business and government agencies.
2. Ancevich, Singrida S., and Payne, R. W. "The Investigation and Treatment of a Reading Disability in a Child of Normal Intelligence," Journal of Clinical Psychology, XVII (October 1961), 416-20.
   Effect of remedial instruction involving learning a new method of writing upon a nine-year-old boy, left-handed but of normal intelligence, with multiple behavior and language difficulties in reading, writing and spelling.
3. Aronow, Miriam S. "A Study of the Effect of Individualized Reading on Children's Reading Test Scores," Reading Teacher, XIV (November 1961), 86-91.
   Comparative study of a group of 351 sixth-grade pupils who had received individualized reading instruction in grades 4 and 5 in New York City with a group of 6,279 sixth graders who had not received such instruction, with respect to their Metropolitan Reading Test scores in grades 3 and 6.
4. Baker, Dean, and MacDonald, James. "Newspaper Editorial Readership and Length of Editorials," Journalism Quarterly, XXXVIII (Autumn 1961), 473-9.
   Analysis of data from The Continuing Study of Newspaper Reading, 1939-1953, to determine the relationship between editorials and selected variables, including editorial length and the number of persons attending to editorials, with respect to the reading of 130 different daily newspapers throughout the country.
5. Belden, Bernard R. "Readability of Biology Textbooks and the Reading Ability of Biology Students," School Science and Mathematics, LXI (December 1961), 689-93.
   Analysis of readability measured by the Dale-Chall formula of five biology textbooks in relation to the reading ability of 357 tenth-grade biology students to determine the potential effectiveness of the texts.
6. Belson, William A. "The Effects of Television on the Reading and the Buying of Newspapers and Magazines," Public Opinion Quarterly, XXV (Fall 1961), 366-81.
   Description of the effects of television viewing on the reading of 53 newspapers and magazines by 1,157 adults in London.
7. Bentall, Grace. "Determination of Critical Levels of Reading Ability for High School Students as Measured by Course Marks and Achievement Test Results," Dissertation Abstracts, XXII (November 1961), 158-9. (University of Oregon).
   Investigation of the relationships between reading level, school achievement in several areas, intelligence, and related course marks for 830 ninth-grade and 720 eleventh-grade randomly selected students.
8. Blakely, W. Paul, and Shadle, Erma M. "A Study of Two Readiness for Reading Programs in Kindergarten," Elementary English, XXXVIII (November 1961), 503-5.
   Comparative study of two groups of 28 kindergarten children with respect to reading readiness following a basal reading program or an activity program of experiences.
9. Bliesmer, Emery P. "Evaluating Progress in Remedial Reading Programs," Reading Teacher, XV (March 1962), 344-50.
   Study of reading progress measured during a nine-month period for groups of 40 and 41 children in grades 4-8 enrolled in remedial reading programs at two reading clinics.
10. Bloomer, Richard H. "The Cloze Procedure as a Remedial Reading Exercise," Journal of Developmental Reading, V (Spring 1962), 173-81.
   Study of improvement in reading

achievement related to the use of the cloze procedure in experimental and control groups of 44 college students enrolled in a remedial reading class.

11. Bloomer, Richard H. “Connotative Meaning and the Reading and Spelling Difficulty of Words,” Journal of Educational Research, LV (November 1961), 107-12.

    Study of the relationship between the emotional intensity and tone of words, and their reading and spelling difficulty, for 149 words judged by college juniors.

12. Bloomer, Richard H. “The Effects of a College Reading Program on a Random Sample of Education Freshmen,” Journal of Developmental Reading, V (Winter 1962), 110-8.

    Study of effects among 40 experimental and 39 control subjects of a 12-week college reading improvement program in terms of reading gains and the correlation of these gains with academic achievement.

13. Boag, Audrey K., and Neild, Margaret. “The Influence of the Time Factor on the Scores of the Triggs Diagnostic Reading Test as Reflected in the Performance of Secondary School Pupils Grouped According to Ability,” Journal of Educational Research, LV (December 1961), 181-3.

    Comparative study of scores for timed and untimed performance on a diagnostic reading test for 722 students in grade 9-12 English classes.

14. Bollenbacher, Joan. “A Study of the Effect of Mobility on Reading Achievement,” Reading Teacher, XV (March 1962), 356-60.

    Report on reading achievement and school mobility of 2,778 boys and 2,800 girls from the 6th grade in Cincinnati in 1960 in comparison with 1952 data.

15. Brown, Charles M. “Whither Basal Reading?” Education, LXXXII (September 1961), 3-5.

    Analysis of mean test results on standardized reading and intelligence tests for 1,500 pupils in grades 3, 5 and 7 who followed a basal reading program supplemented by independent reading.

16. Brueckman, Elizabeth A. “Advanced Reading Program-Taft High School,” Journal of Developmental Reading, V (Summer 1961), 228-38.

    Tabulation of test scores according to scholastic aptitude on standardized tests given in connection with a reading program for 96 high school juniors and seniors.

17. Calvin, Allen D., and Hanley, Charles. “Exposure to Programmed Learning Material and Reading Comprehension,” Psychological Reports, X (April 1962), 538.

    Study of effects of programmed instructional material on reading rate and level of comprehension for an experimental group of 215 students and a control group of 156 students enrolled in high school mathematics courses.

18. Carner, Richard L. “The Effect of Televised Reading Instruction on Attitudes Toward Reading,” Elementary English, XXXIX (March 1962), 234-6.

    Questionnaire survey regarding change in attitude following televised reading instruction among 111 fifth- and sixth-grade pupils, analyzed in relation to level of reading performance.

19. Cleland, Donald T., and Toussaint, Isabella. “The Interrelationship of Reading, Listening, Arithmetic Computation, and Intelligence,” Reading Teacher, XV (January 1962), 228-31.

    Correlation analysis of test scores on reading, listening, arithmetic computation, and two tests of scholastic ability for 172 pupils in grades 4, 5, and 6.

20. Cohan, Mayme. “Two and a Half and Reading,” Elementary English, XXXVIII (November 1961), 506-8.

    Report of an attempt to teach a two- and a half-year-old girl to read isolated words and sentences.

21. Coleman, E. B. “Improving Comprehensibility by Shortening Sentences,” Journal of Applied Psychology, XLVI (April 1962), 131-4.

    Study of the comparative readability measured by the cloze procedure for 90 college undergraduates of three versions of the same difficult material written with average sentence lengths of 15, 23, and 39 words.

22. Collins, Ray. “The Comprehension of Prose Materials by College Freshmen When Read Silently and When Read Aloud,” Journal of Educational Research, LV (October 1961), 79-82.

    Controlled experimental study of comprehension scores for 60 college freshmen following the oral and silent reading of seven stories of varying levels of difficulty according to the Flesch readability formula.

23. Colvin, Charles. “ ‘Ideal’ College Reading Program,” Journal of Developmental Reading, V (Winter 1962), 77-81.

    Formulation of an “ideal” college reading program from detailed questionnaires returned by representatives of 42 colleges and universities in Pennsylvania.

24. Curry, Robert L., and Hughes, Hughie. “Relationship Between Measured and Anticipated Achievement in Reading,” Junior College Journal, XXXII (October 1961), 91-6.

    Study of relationships between anticipated and measured achievement in reading vocabulary and comprehension for 640 high school sophomores classified into three groups according to intellectual ability.

25. DeCharms, Richard, and Moeller, Gerald. "Values Expressed in American Children's Readers: 1800-1950," *Journal of Abnormal and Social Psychology*, LXIV (February 1962), 136-42.

    Analysis of the incidence of achievement and affiliation imagery, and of moral teaching, in a systematic sampling of children's readers at 20-year intervals from 1800-1950.

26. Docter, Robert L. "Reading Workbooks: Boon or Busywork?" *Elementary English*, XXXIX (March 1962), 224-8, 230.

    Examination of the effectiveness of reading workbooks in 36 classrooms in 18 randomly selected elementary schools in Los Angeles, including questionnaire reactions from 1,200 teachers and 80 administrators concerning their use.

27. Durkin, Dolores. "Children Who Learned to Read at Home," *Elementary School Journal*, LXII (October 1961), 15-8.

    Comparison of reading grade levels over a two-year period with I. Q. and other early characteristics of 49 children who learned to read prior to entering school.

28. Educational Records Bulletin No. 79. *1961 Achievement Testing Program in Independent Schools and Supplementary Studies, July 1961*. New York: Educational Records Bureau. Pp. 84.

    Annual summary and correlation analysis of a comprehensive public and private school achievement testing program, including reading achievement, for elementary and secondary schools.

29. Educational Records Bulletin No. 80. *1961 Fall Testing Program in Independent Schools and Supplementary Studies, February 1962*. New York: Educational Records Bureau. Pp. 69.

    See preceding item.

30. Ewalt, H. Ward. "Visual Performance and Its Relationship to Reading Achievement," *Journal of the American Optometric Association*, XXXIII (June 1962), 825-9.

    Examination of functional visual performance of 50 retarded readers, 84 percent having 20/30 or better vision, and all of whom passed the Snellen test for distance vision.

31. Feinberg, Mortimer R., Long, L., and Rosenheck, V. "Results of a Mandatory Study Course for Entering Freshmen," *Journal of Developmental Reading*, V (Winter 1962), 95-100.

    Comparison of mean reading achievement in a mandatory reading and study skills program for a group of 64 college freshmen in 1960 in relation to a matched group not receiving such a program in 1959.

32. Felstehausen, Herman. "Headline Motivation and Reader Response in a Farm Magazine," *Journalism Quarterly*, XXXIX (Spring 1962), 216-9.

    Study of the effect of headlines differing in motivational appeal upon readership of agricultural news articles as determined by interviews with 200 male and 201 female readers.

33. Fields, Irwin H. "The Centinela Valley Plan: A Mandatory Reading and Study Skills Program for Ninth Grade Students," *Journal of Developmental Reading*, IV (Summer 1961), 254-9.

    Report of improvement on reading test scores in connection with a six-week reading program given to freshmen as part of English in four high schools.

34. Fisk, Robert R. "Leisure Reading in Alberta," *Alberta Journal of Educational Research*, VII (September 1961), 134-9.

    Questionnaire survey of 11,000 Canadian students in junior and senior high schools and 2,000 teachers regarding reading interests, habits, instruction, facilities, and attitudes.

35. Garwood, Victor P., and Boyer, Ernest L. "Effects of Delayed Sidetone and Speech Content on Elapsed Reading Time," *Journal of Communication*, XII (March 1962), 44-50.

    Experimental study among 30 college undergraduates of the length of oral reading time for three types of emotionally toned sentences under conditions of delayed and non-delayed sidetone feedback.

36. Gates, Arthur I. "Vocabulary Control in Basal Reading Material," *Reading Teacher*, XIV (November 1961), 81-5.

    Report of two studies, involving a total of 412 pupils in grades 2 and 3, of the pupils' ability to derive the correct meaning of new words not yet encountered in their basal readers.

37. Gates, Arthur I. "The Word Recognition Ability and the Reading Vocabulary of Second and Third Grade Children," *Reading Teacher*, XV (May 1962), 443-8.

    Study of the efficiency of two classes of second-grade pupils in recognizing new words from third- and fourth-grade basal readers.

38. Geake, Robert R. "Michigan Schools Stress Special Reading Programs," *Michigan Education Journal*, XXXIX (November 1961), 262-3.

    Questionnaire survey of a systematic sample of Michigan high schools to determine existing and proposed special reading programs.

39. Greenberg, Bradley, and Tannenbaum, Percy. "The Effects of Bylines on Attitude Change," Journalism Quarterly, XXXVIII (Autumn 1961), 535-7.

Pre- and post-effects of the presence, placement or absence of prestige bylines on attitude changes determined by semantic differential analysis for 216 university undergraduates in reading persuasive presentations about three topics.

40. Grissom, Loren. "Characteristics of Successful Reading Improvement Programs," English Journal, L (October 1961), 461-4.

Study of 15 reading improvement programs in Illinois, following a state-wide survey, to obtain further information about the characteristics of the programs through staff discussions and class observations.

41. Groff, Patrick J. "Children's Attitudes Toward Reading and Their Critical Reading Abilities in Four Content-Type Materials," Journal of Educational Research, LV (April 1962), 313-8.

Correlation of pupil attitudes and scores on an experimental critical reading test for 142 boys and 163 girls in the fifth and sixth grades.

42. Groff, Patrick J. "Prediction of the Reading Difficulty of Children's Magazines," California Journal of Educational Research, XIII (January 1962), 39-45.

Application of the Dale-Chall readability formula to samples of two issues, six months apart, of 14 children's magazines.

43. Groff, Patrick J. "A Survey of Basal Reading Group Practices," Reading Teacher, XV (January 1962), 232-5.

Compilation of reports from student teachers regarding the amount and time of shifting of pupils among basal reading groups for 2,179 second-, third-, and fourth-grade pupils in 66 classes.

44. Gruber, Ellis. "Reading Ability, Binocular Coordination and the Ophtholmograph," Archives of Ophthamology, LXVII (March 1962), 280-8.

Description of orthoptic analysis and ophthalmograph findings on binocular coordination for 50 patients, aged 10-68 years, selected at random from private practice.

45. Harcum, E. R., and Dyer, D. W. "Monocular and Binocular Reproduction of Binary Stimuli Appearing Right and Left of Fixation," American Journal of Psychology, LXXXV (March 1962), 56-65.

Tachistoscopic measurement of monocular and binocular accuracy for 12 adults in perceiving targets to the right and left of a fixation point.

46. Harris, Theodore L. "Summary of Investigations Relating to Reading, July 1, 1960 to June 30, 1961," Journal of Educational Research, LV (February 1962), 197-220.

Summary and bibliographical notations on 148 investigations in reading appearing in 1960 and 1961.

47. Hess, Robert D., and Goldman, Harriet. "Parents' Views of the Effect of Television on Their Children," Child Development, XXXIII (June 1962), 411-26.

Analysis of interview data for 99 mothers in a metropolitan area concerning a wide range of effects of television viewing on their children.

48. Hinton, Evelyn. "Dropout Rate and Academic Progress of Two Groups of Students Enrolled at the University of Wichita," Journal of Developmental Reading, IV (Summer 1961), 272-5.

Report of the drop-out rate and scholastic averages of 71 college students in a non-credit reading improvement course and 24 honors section students.

49. Ibeling, Frederick W. "Supplementary Phonics Instruction and Reading and Spelling Ability," Elementary School Journal, LXI (December 1961), 152-6.

Comparison of reading and spelling test scores for a control and experimental group receiving additional phonics instruction among 600 second-, fourth- and sixth-grade pupils.

50. Irwin, Orvis C. "Infant Speech: Effect of Systematic Reading of Stories," Journal of Speech and Hearing Research, III (June 1960), 187-90.

Experimental study over a 17-month period of phonetic production in infants as related to exposure to systematic reading of stories, using an experimental group of 24 and a control group of 10 infants from working-class homes.

51. Johnson, Lovell R., and Platts, D. "A Summary of a Study of the Reading Ages of Children Who Had Been Given Remedial Teaching," British Journal of Educational Psychology, XXXII (February 1962), 66-71.

Comparison of reading scores obtained one year after remedial teaching, and correlations of these with several variables from prior data, for 74 children receiving individual remedial reading instruction and 210 receiving group instruction over a five-year period in England.

52. Jones, Reginald L., and Van Why, Earl L. "The SRA Reading Laboratory and Fourth-Grade Pupils," Journal of Developmental Reading, IV (Autumn 1961), 36-46.

Comparison of mean vocabulary and reading comprehension scores for 104 fourth-grade children divided into equal control and experimental groups which used differentiated reading materials.

53. Keating, Leslie E. "A Pilot Experiment in Remedial Reading at the Hospital School, Lingfield, 1957-60," British Journal of Educational Psychology, XXXII (February 1962), 62-5.

Report on effect of group and individualized remedial reading instruction upon seriously retarded boys, aged 11-4 to 16-2 years, in an English hospital school.

54. Kelly, Thomas. "Utilization of Filmstrips as an Aid in Teaching Beginning Reading," Dissertation Abstracts, XXII (September 1961), 817. (Indiana University).

Comparative study of experimental and control classes, involving 138 children in the second semester of grade 1, in using filmstrips to present basal text materials and lessons in phonetic and structural analysis from the teacher's manual.

55. Kermoian, Samuel B. "Teacher Appraisal of First Grade Readiness," Elementary English, XXXIX (March 1962), 196-201.

Correlations between teacher judgments on reading readiness and test scores on the Metropolitan Readiness Test for 276 first-grade children.

56. Keshian, Jerry G. "How Many Children are Successful Readers?" Elementary English, IX (October 1961), 408-10.

Comparison of scores on standardized group tests of reading and mental ability group tests for 406 fifth-grade pupils drawn from three communities differing in socio-economic level, including further examination of the characteristics of 72 successful readers.

57. Lasater, G. M. "Reading Epilepsy," Archives of Neurology, VI (June 1962), 492-5.

Clinical report of four adult cases of epileptic seizure which occur only during the reading act.

58. Lawson, John R., and Airla, Donald. "Comparison of Wide Range Achievement Test and Gray Oral Reading Paragraphs Reading Scores of Mentally Retarded Adults," Perceptual and Motor Skills, XIV (June 1962), 474.

Correlation of test scores on two reading achievement tests for 30 mentally retarded adults.

59. Love, Harold. "An Experimental Phonics Program Versus a Controlled Integral Reading Program," Journal of Developmental Reading, IV(Summer 1961), 280-2.

Comparison of progress in a remedial reading program for two groups of 15 children matched for age, intelligence and reading ability, who were given phonics instruction as separate drill in workbooks vs. phonics as a part of a broader reading program over a six-week period.

60. Lubershane, Melvin. "Can Training in Listening Improve Reading Ability?" Chicago Schools Journal, XLIII (March 1962), 277-81.

Comparison of reading growth for 72 fifth-grade pupils, 35 of whom received exercises in responding to oral instructions, between reading improvement and auditory training.

61. Lyle, Jack. "Immediate vs. Delayed Reward Use of Newspapers by Adolescents," Journalism Quarterly, XXXIX (Winter 1962), 83-5.

Questionnaire study of newspaper reading habits of 674 sixth graders and 700 tenth graders in five Rocky Mountain towns in reading for immediate reward (entertainment) vs. delayed reward (news and editorials).

62. McCracken, Robert A. "The Oral Reading Performance of a Second Grade Class Using an Informal Reading Test," Journal of Educational Research, LV (November 1961), 113-7.

Analysis of reading time, comprehension, and number and type of errors made in oral reading of four graded paragraphs by 36 second-grade pupils classed as good, average, or poor readers.

63. McCracken, Robert A. "Standardized Reading Tests and Informal Reading Inventories," Education, LXXXII (February 1962), 366-9.

Comparison of reading comprehension and vocabulary achievement test scores with the results of an informal reading inventory for 56 sixth-grade pupils.

64. McDonnell, M. W. "The Prediction of Academic Achievement of Superior Grade Three Pupils," Alberta Journal of Educational Research, VIII (June 1962), 111-8.

Correlation study of three group intelligence tests with the total score and four sub-scores of the primary battery of California Achievement Tests for 178 third-grade Canadian pupils having scores of 120 or above on the California Test of Mental Maturity.

65. McElroy, Kathryn M. "A Comparative Study of Reading Growth from Grades Two to Seven," Reading Teacher, XV (November 1961), 98-101.

Discussion of trends in median vocabulary and reading comprehension scores for an unspecified number of pupils of above average ability in grades 2-7 in an upper-class community.

66. McGinnis, Dorothy J. "The Preparation and Responsibility of Secondary Teachers in the Field of Reading," Reading Teacher, XIV (November 1961), 92-7.

Questionnaire reports from 570 secondary teachers and 1,029 college freshmen on various aspects of secondary reading instruction.

67. McLeod, John. "The Estimation of Readability of Books of Low Difficulty," *British Journal of Educational Psychology*, XXXII (June 1962), 112-8.

    Report of two experiments involving 255 and 375 children, respectively, in grades 2 to 4 to estimate the reading difficulty of books 1, 2, 3, and 5 of a reading series, the *Wide Range Readers*, according to the known reading achievement levels of pupils.

68. McNelly, John T. "Meaning Intensity and Interest in Foreign News Topics," *Journalism Quarterly*, XXXIX (Spring 1962), 161-8.

    Application of semantic differential scale analysis to a study of the relationship of meaning intensity to reader interest in foreign news topics for 104 high school and 126 college students.

69. Magoon, Maxwell, and Magoon, Thomas M. "A Description of the University of Maryland's Reading and Study Skills Laboratory," *Journal of Developmental Reading*, V (Spring 1962), 182-8.

    Description of academic abilities of students in a college reading laboratory in 1959-60, including data on progress in reading speed and comprehension for college students as a function of the number of reading laboratory sessions attended.

70. Manis, Melvin. "The Interpretation of Opinion Statements as a Function of Message Ambiguity and Recipient Attitude," *Journal of Abnormal and Social Psychology*, LXIII (July 1961), 76-81.

    Investigation of attitudes expressed toward fraternities by 66 male undergraduates and their interpretation of ambiguous statements about fraternities.

71. Manis, Melvin. "The Interpretation of Opinion Statements as a Function of Recipient Attitude and Source Prestige," *Journal of Abnormal and Social Psychology*, LXIII (July 1961), 82-6.

    Investigation of attitudes expressed by 65 male undergraduates toward fraternities and statements reflecting high and low prestige toward fraternities.

72. Martin, Marvin. "Fifty Books They Can't Resist," *Elementary English*, XXXIX (May 1962), 415-7.

    Compilation of a favorite book list based upon recommendations of the three best books read by each of 250 sixth-grade students in an Illinois suburban community.

73. Maw, Wallace H., and Maw, Ethel W. "Children's Curiosity as an Aspect of Reading Comprehension," *Reading Teacher*, XV (January 1962), 236-40.

    Report of two studies of 102 and 458 fifth-grade pupils ranked by teachers, peers, and selves as high and low in curiosity in relation to their performance on a "Foolish Sayings Test."

74. Mingoia, Edwin. "Possible Causes of Underachievement in Reading," *Elementary English*, XXXIX (March 1962), 220-3.

    Analysis of pupil and environmental characteristics associated with underachievement and overachievement in reading, based upon three studies involving over 3,000 fourth, fifth and sixth graders.

75. Money, John, ed. *Reading Disability: Progress and Research Needs in Dyslexia.* Baltimore: Johns Hopkins Press, 1962. Pp. 222.

    Presentation and summary of papers given at the Johns Hopkins Conference on Research Needs and Prospects in Dyslexia and Related Aphasic Disorders in 1961.

76. Muehl, Siegmar. "The Effects of Visual Discrimination Pretraining with Word and Letter Stimuli on Learning to Read a Word List in Kindergarten Children," *Journal of Educational Psychology*, LII (August 1961), 215-21.

    Controlled discrimination training experiment with 60 kindergarten pupils in learning to recognize three words as a result of differential pretraining in nonsense word sets.

77. Newton, James E. "Personal Preference for Method as a Factor in Learning," *Journal of Educational Psychology*, LIII (February 1962), 43-7.

    Experimental study of effect upon achievement of reading versus lecture presentation, including interactions among preferred, not preferred, or no preference for method of presentation for groups of airmen in basic training.

78. North, Marie. "Measurable Gains Made by High School Students in a Developmental Reading Course," *Journal of Developmental Reading*, V (Spring 1962), 208-13.

    Report of progress made by 368 high-school seniors receiving developmental reading training at the beginning and end of a semester.

79. Nyman, Patricia, Kearl, Bryant E., and Powers, Richard D. "An Attempt to Shorten the Word Test with the Dale-Chall Readability Formula," *Educational Research Bulletin*, XL (September 1961), 150-2

    Report of predictive power and error of the Dale-Chall readability formula when the basic word list was shortened from 3000 to 920 words.

80. O'Donnell, Roy. "Awareness of Grammatical

Structure and Reading Comprehension," High School Journal, XLV (February 1962), 184-8.

Correlation study of standardized test scores for vocabulary, reading comprehension, grammar, and for scores on a specially constructed test of awareness of sentence structure among 101 high-school seniors.

81. Otto, Wayne. "The Acquisition and Retention of Paired Associates by Good, Average, and Poor Readers," Journal of Educational Psychology, LII (October 1961), 241-8.

Controlled experimental study of acquisition and retention of paired associates (geometric forms and trigrams) using three modes of presentation in relation to good, average, and poor reading test performance by 108 children from grades 2, 4 and 6.

82. Otto, Wayne, and Lasswell, Ann. "Relationship of Secondary Beginning Strokes in Handwriting to Reading Ability," Perceptual and Motor Skills, XIV (June 1962), 530.

Analysis of differences between 25 good readers and 25 poor readers of normal intelligence in grades 7 and 8 in the use of a secondary beginning stroke in handwriting, considered as a symptom of immature personality development.

83. Peterson, Wilbur, and Thorp, Robert. "Weeklies' Editorial Effort Less than 30 Years Ago," Journalism Quarterly, XXXIX (Winter 1962), 53-6.

Analysis of proportion of editorial content of 215 non-daily newspapers in Iowa in 1960 in comparison to similar data for 1930.

84. Ramsay, Wallace. "A Study of Salient Characteristics of Pupils of High and Low Reading Ability," Journal of Developmental Reading, V (Winter 1962), 187-94.

Comparison of questionnaire reports from 61 poor readers and 81 good readers in grades 4, 5, and 6 regarding their reading interests, habits and self-concepts as readers.

85. Rothkoff, Ernest Z. "Learning from Written Sentences," Psychological Reports, X (June 1962), 667-74.

Study of the retention effects of the order of presentation of topically related but loosely connected sentences, and the location of information within sentences, for 114 female switchboard operators.

86. Russell, David H. "An Evaluation of Some Easy-to-Read Trade Books for Children," Elementary English, XXXVIII (November 1961), 475-82.

Application of several modes of evaluation to 10 "easy-to-read" trade books for children.

87. Sampson, Olive C. "Reading Skill at Eight Years in Relation to Speech and Other Factors," British Journal of Educational Psychology, XXXII (February 1962), 12-7.

Study of the relationship between speech and reading development, with additional reference to personal adjustment for 48 eight-year-old children.

88. Schrock, Ralph E., and Grossman, Milton. "Pilot Study: Motivation in Reading," Reading Teacher, XIV (November 1961), 119-21.

Study of effect of visual training and encouragement on oral reading performance of experimental and control groups of seventh-grade students with normal I.Q. but reading two years below expectancy.

89. Schwartz, Martin. "The Effect of Delayed Sidetone Upon the Reading Rate of Whispered Speech," Speech Monographs, XXVIII (November 1961), 293-5.

Analysis of variance of the whispered reading time for reproducing 15 five-syllable phrases under different degrees of delayed sidetone feedback by 10 male college students.

90. Shabaglian, John A. "A Study of the Effects of Deanol, a Psychic Energizer, on the Learning of Reading by a Group of Mentally Retarded Boys, Ages 12 through 15," National Association of Secondary Schools Principals Bulletin, XLVI (February 1962), 174-5.

Experimental study of the effect of Deanol on the quiz scores and reading test scores of 30 mentally retarded boys with I.Q. scores of 50-69.

91. Siegel, Max. "Adult Reading Improvement: A Five Year Report," Reading Teacher, XV (January 1962), 246-53.

Analysis of population characteristics and results of an adult reading program conducted over a five-year period with local college staff collaboration in a library setting.

92. Smith, Ruth. "Children's Reading Choices and Basic Reader Content," Elementary English, XXXIX (March 1962), 202-9.

Comparison of reading interests, classified in 23 categories, of 110 first-grade pupils expressed through choice of library books with interest categories found in 37 pre-primers and primers.

93. Sommers, Ronald R., et al. "Effects of Various Durations of Speech Improvement Upon Articulation and Reading," Journal of Speech and Hearing Disorders, XXVII (February 1962), 54-61.

Analysis of variance of the effects of varying lengths of speech improvement classes upon reading improvement for large groups of first- and second-grade pupils.

94. Sorelle, Zell, and Walker, Jack. "What Is

Television Doing to Our Children?" The Journal of Educational Research, LV (February 1962), 236-7.

Questionnaire survey of the television viewing habits of 2,000 students when they were in junior high school and three years later when they were in senior high.

95. Spache, George D., and Tillman, Chester E. "A Comparison of the Visual Profiles of Retarded and Non-Retarded Readers," Journal of Developmental Reading, V (Winter 1962), 101-9.

Comparison of differences in the means and the percent of failure on Ortho-Rater visual test scores for 114 retarded readers and 101 normal readers of comparable mental age in the elementary school.

96. Staiger, Ralph C. "Medicine for Reading Improvement," Journal of Developmental Reading, IV (Autumn 1961), 48-51.

Analysis of variance of reading test scores as influenced by Deanol for 60 matched pairs of retarded readers drawn from elementary, high school and college populations.

97. Stanchfield, Jo M. "Reading Interests of Eighth-Grade Boys," Journal of Developmental Reading, V (Summer 1962), 256-65.

Report on reading interests identified by interviews and rating scales for 51 eighth-grade boys selected by stratified random sampling.

98. Standards of Reading, 1948 to 1956. Pamphlet No. 32 of the British Ministry of Education. London: Her Majesty's Stationery Office, 1957. Pp. 46.

Comparative surveys of the reading achievement of 11- and 15-year-old children in England in 1952 and 1956 in relation to a previously reported 1948 survey of England and Wales.

99. Stempel, Guido H. "Content Patterns of Small and Metropolitan Dailies," Journalism Quarterly, XXXIX (Winter 1962), 88-90.

Comparative analysis of the content of four largest and four smallest daily newspapers of a midwestern state for a one-week period.

100. Stempel, Guido H. "Selectivity in Readership of Political News," Public Opinion Quarterly, XXV (Fall 1961), 400-4.

Results of interviews with 205 college students regarding the reading of circulated articles on campus candidates.

101. Stone, David R. "Speed of Idea Collecting in University Study-Reading," Journal of Developmental Reading, V (Spring 1962), 149-56.

Report on progress in and correlation among different modes of speed reading attack for 62 college students.

102. Strickland, Ruth. "The Language of Elementary School Children: Its Relationship to the Language of Reading Textbooks and the Quality of Reading of Selected Children," Bulletin of the School of Education, Indiana University, Vol. XXXVIII, No. 4, July 1962.

Analysis of oral language patterns of a random sampling of 575 pupils in grades 1-6 in an Indiana city in relation to language patterns in selected basal reading textbooks and to other within-subject variables.

103. Tannenbaum, Percy, and Lynch, Mervin. "Sensationalism: Some Objective Message Correlates," Journalism Quarterly, XXXIX (Summer 1962), 317-23.

Analysis of stylistic attributes correlated with degree of sensationalism in passages read by 61 adults.

104. Torrance, E. Paul, and Harmon, Judson. "Effects of Memory, Evaluative, and Creative Reading Sets on Test Performance," Journal of Educational Psychology, LII (August 1961), 207-14.

Study of the effects of three textbook reading sets upon the test performance of 115 college graduate students, measured at the end of each of three one-week periods.

105. Troxel, Vernon. "The Effects of Purpose on the Reading of Expository Mathematical Materials in Grade Eight," The Journal of Educational Research, LV (February 1962), 221-7.

Correlations between reading speed, reading accuracy, general reading ability, and arithmetic achievement scores for two groups, totaling 45 eighth-grade students, who read the same expository mathematics materials for different purposes.

106. Vernon, Philip E. "The Determinants of Reading Comprehension," Educational and Psychological Measurement, XXII (Summer 1962), 269-86.

Study of intercorrelations and factors involved in testing vocabulary and reading comprehension in several ways, by essay vs. multiple-choice methods, among 108 male English college students and 75 American college students of both sexes.

107. Walters, Richard H., Loan, Malle Van, and Crofts, Irene. "A Study of Reading Disability," Journal of Consulting Psychology, XVII (August 1961), 277-83.

Comparison of group test scores on selected perceptual and personality variables for 58 good, average and poor male readers in grades 3-6.

108. Warren, Mary Bay. "The Massapiqua Junior High School Reading Program," Journal of Developmental Reading, V (Summer

1962), 245-55.

Experimental study on change in test performance for 50 eighth-grade students, half of whom followed a reading program using reading instrumentation and half of whom did not.

109. Weaver, C.H., Furbee, Catherine, and Everhart, Rodney. "Articulatory Competency and Reading Readiness," Journal of Speech and Hearing Research, III (June 1960), 174-80.

Comparison of Gates Readiness Test scores and articulation errors among 638 first-grade children.

110. Wilcox, Walter. "The Press of the Radical Right: An Exploratory Analysis," Journalism Quarterly, XXXIX (Spring 1962), 152-60.

Description of thematic patterns and the position of the radical rightist press on the political opinion spectrum, based upon an analysis of an extensive sample of 42 periodicals and a more intensive study of randomly drawn issues from two periodicals.

111. Wood, Roger L. "Attrition as a Criterion for Evaluation of Non-Credit College Reading Programs," Journal of Developmental Reading, IV (Autumn 1961), 27-35.

Study of the extent to which personality and selected instructional variables appear to affect attrition in a non-credit college reading improvement program, based upon a representative sample of 701 students meeting in classes for 6 or 7 weeks.

112. Wyatt, Nita. "A Study of the Relationship of Extensive Reading to Certain Writing Skills of a Selected Group of Sixth Grade Children," University of Kansas Bulletin on Education, XVI (November 1961), 13-8.

Investigation of the relationship between amount of voluntary reading and various facets of writing as expressed in four original compositions for 65 children with C. A.'s between 10-9 and 11-9 and grouped according to IQ.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 6, February 1963)

A Study of Experimental Methods for Teaching Emotionally Disturbed, Brain Damaged, Retarded Readers[*1]

MAX TALMADGE, ANTHONY DAVIDS, MAURICE W. LAUFER
Emma Pendleton Bradley Hospital
Riverside, Rhode Island

IN RECENT YEARS there has been a growing awareness that there is a distinct syndrome frequently associated with children who are slow learners, especially retarded readers, which has been characterized under such headings as "hyperkinetic disorder" (Laufer, Denhoff and Solomons, 1957), "minimally brain damaged" (Burks, 1957), "visual-motor impairment" (Strauss and Lehtinen, 1947), and "primary reading disturbance" (Rabinovitch et al., 1956). Generally, these children are thought to have some central nervous system abnormality of function, subcortical or cortical or both, which has interfered with their ability to receive or retain and reproduce visual, auditory or motor cues, as well as inducing the behavioral syndrome. Behaviorally, these children can often be identified in that they have a short attention span, are hyperactive, easily distracted, and readily overwhelmed by anything new or different. They reveal tendencies to perseverate, with considerable difficulty in shifting, and poor ability to abstract and to generalize. In reading they often reverse words (such as "saw" for "was"), are unable to pronounce words and fail to develop efficient techniques in reading new words. Their spelling is usually quite poor because they have no understanding of letter sounds and have a poor concept of phonics; there may be mirror writing without awareness and their handwriting is often hardly legible. These symptoms have been given detailed attention by Strauss and Lehtinen (1947) and Strauss and Kephart (1955), and descriptions of these children can be found in recent articles by Laufer et al., (1957) and Rabinovitch et al., (1956).

The purpose of the present paper is to describe a remedial educational approach that has been found useful with children who evidence such difficulties. This experimentation has been conducted at the Emma Pendleton Bradley Hospital, which is a residential treatment center for emotionally disturbed children who have the potential of at least normal intelligence.

The Experiment

Procedure and Subjects

The California Achievement Battery was administered during the school year 1957-58 and it was found that 24 of the 42 children tested were at least two years retarded in reading. These 24 severely retarded readers were evaluated for evidence of brain damage. This was done by a neuro-psychiatrist who utilized his clinical judgment based upon the case history, neurological examination, EEG recordings, and photo-metrazol tests (Laufer, 1957). On this basis, eight children were thought to have cortical dysfunction, six fell into the questionable category where cortical impairment was suspected but not definitely indicated, and ten children revealed no indication of cortical dysfunction. These children were divided into an experimental and control group so that an equal number of each category were in both groups. The selection of children for the two groups depended solely upon obtaining equated groups for such variables as age, intelligence, and reading retardation.

As shown in Table I, the differences between the groups were very slight and well matched groups resulted. The children in the experimental group all received kinesthetic and auditory approach with its many special aids and novel techniques (e.g. Durrell, 1940, Fernald, 1943; French, 1953). The children in the control group received the traditional method. All children were alike in being emotionally disturbed. The children in each group were taught separately although many of the teachers had both experimental and control children, as well as non-retarded children, at different times in separate classes. Thus, most of the teachers had some experimental and control children, reducing the possibility that the "teacher" factor might be a source of contamination. The program started on March 1, 1958 and continued for three months until the end of the school year. Approximately one

---

[*] Footnotes will be found at end of article

class hour per day was devoted to the reading program for all children included in this study. Retesting was conducted during the month of June.

Summary of Method

An outline and description of the kinesthetic aids is contained in the outline placed after the summary at the end of this paper. Here we will present a brief summary of the methods and approach that were utilized in this study.[2]

First, activities were provided for recognition of differences and likenesses in visual configuration patterns, beginning with objects and proceeding to letter forms. For example, they were shown series of letters, such as p d b, and were asked to point out similarities and differences. Instruction in the recognition of different vowel sounds was then added. For example, the letter "A" was printed on the blackboard, using chalk of a distinctive color that would be consistently associated with this vowel, and examples were given of how this vowel would sound differently when contained in different words. Pages were selected from magazines and the children were asked to circle all the "A" using a pencil of the color to be associated with this vowel. Pictures of interesting objects (e. g. elephant, Indian, etc.) were used to aid the children in remembering the various vowel sounds that appeared in these words. Short sentences were presented, emphasizing the sound of the vowel each time it appeared.

Much use was made of block letters specially prepared for this project. These consisted of large rectangles, deep enough for a letter to be cut out of the block of wood, of a three-dimensional nature, which could be handled and seen in all three dimensions. Again, each vowel had its own color, and the block for a particular vowel always appeared in its special color (e. g., red for "a"). The letter cut out of the blocks, though of the same color, also had sandpaper grit sprinkled upon it. Thus, it was slightly lighter in color than its surroundings when put back into the block from which it had been cut, and now appeared in the two dimensions characteristic of printed material. The sandpaper grit, however, offered resistance when the finger was rubbed along the outline of the letter, thereby allowing the aid of another sense modality in the child's attempt to learn to correctly recognize the given letter.

Another helpful technique involved the tracing of letters by the forefinger in wet sand. That is, utilizing a large tray containing wet sand, the child traced letters and printed words with his finger, employing gross body movements in overcoming the resistance offered by the medium. While tracing the letters and words in the sand, the child simultaneously pronounced them aloud.

The letters of each vowel for writing manuscript (not cursive writing) were then presented. Only vowels were presented, in isolation, and not the consonants. The basic vocabulary of the first preprimer was then started and discrimination of letters and word forms were combined. Visual and auditory discriminations were taught concurrently, not only for recognition of independent words, but also to acquire the first words of the reading vocabulary.

As the program progressed, words presented previously were reviewed by means of a home made tachistoscope which permitted varying the speed of exposure for words of varying degrees of difficulty. As new words were added, they were carefully analyzed for visual and auditory components. Sentences describing interesting activities and brief stories of experiences they liked were used increasingly as the group finally progressed to reading from the conventional primer.

Results

Objective Findings

In view of the relatively small samples, with so few children in each of the sub-categories, conventional statistical evaluation of the results would not be too meaningful. It seems more appropriate to look for consistencies and noteworthy trends. As can be seen in Table II, there was noticeably greater improvement in the experimental group. The greatest gain was evidenced by children in the organic sub-group who were taught by the experimental method. These children gained a full grade in reading ability in the course of the three-month experimental period. The next largest improvement was found in the experimental sub-group of children who had questionable organic deficit. And even the children in the experimental group who did not show evidence or organic involvement gained more than their control subjects as a result of the experimental treatment. In comparing the groups as a whole, it is noteworthy that the mean improvement score in the experimental group was almost twice as great as the mean improvement in the control group. By way of further comparison, the group of non-retarded readers, who were not actually involved in this research program, gained just about three and a half months in reading ability during the three months interval, which is less than half the mean improvement in the experimental group.

In addition to these encouraging test findings, there appeared to be some connection between improved motivation in the school situation and overall therapeutic improvement. It was noted that the children who made pronounced gains in school achievement also seemed to become more secure within themselves.

In general, then, the findings suggest that the kinesthetic procedure was most beneficial for children with cerebral dysfunction. Moreover, there were indications that emotionally disturbed children who have severe reading difficulties de-

TABLE I

SUBJECTS USED IN EXPERIMENT

Experimental Group					Control Group				
Name	CNS Dysfunction	Age	IQ	Reading Retardation*	Name	CNS Dysfunction	Age	IQ	Reading Retardation*
Gregory	No	8-6	87	2-2	James	No	10-7	93	3-3
Bill	?	12-9	95	3-8	Stuart	Yes	11-11	85	3-3
Ned	Yes	11-1	84	5-2	George	Yes	10-5	82	3-1
Pauline	?	10-11	90	3-1	Michael	No	12-7	73	2-4
Richard	Yes	10-8	93	3-9	Ronald	No	11-4	83	4-6
Charles	?	10-1	78	2-4	Lee	No	12-10	98	2-8
Frank	No	9-6	95	3-2	Tess	Yes	10-4	88	2-1
Patrick	No	9-11	72	2-9	Katherine	?	10-7	86	2-5
Jim	Yes	9-6	83	3-3	Lloyd	?	11-2	104	3-6
Walter	No	9-7	99	2-4	Perry	?	9-3	82	3-4
Donald	No	11-4	109	4-5	William	Yes	9-9	69	3-0
Jeffrey	Yes	11-5	76	3-8	Nicholas	No	9-11	77	2-6
Average =		10-5	88.4	3-3	Average =		10-11	85.0	3-0

*Years of retardation as indicated by reading level obtained on the California Achievement Tests in comparison with level expected on the basis of their chronological ages.

TABLE II

IMPROVEMENT IN READING ACHIEVEMENT SCORES FOLLOWING THREE MONTH EXPERIMENTAL PERIOD

	Mean Improvement Scores on California Achievement Test		
	Vocabulary	Comprehension	Total Reading
Experimental Group			
Organic	.94	1.14	1.04
Non-organic	.64	.62	.62
Questionable	.60	1.03	.74
Total (N = 12)	.73	.93	.80
Control Group			
Organic	.40	.53	.43
Non-organic	.54	.12	.42
Questionable	.20	.77	.46
Total (N = 12)	.38	.47	.44
Non-retarded Group (N = 18)	.24	.46	.36

Note – These scores are in terms of 10-month academic years, with a score of 1.00 indicating one grade level.

rive more benefit from this experimental method than they do from the traditional approach to childhood education. It must be emphasized, however, that the number of children studied was small, there was considerable variabiity within the groups, and the experimental period was quite brief. Thus, this must be viewed as a pilot investigation, with need for further study, on a much larger scale and covering a much longer interval, before these findings can be regarded as conclusive.

Illustrative Case Example

Richard (Age 10-8, WISC IQ 93): Richard was referred to Bradley with a varied list of complaints including overeating, hyperactivity, destructiveness, sleeplessness, aggression towards younger siblings, stealing, and fire setting. Richard began school at Bradley on a pre-primary level in 1954. He was then 7 years of age and was reported as being much more capable than his school work revealed. He followed a consistent negative pattern of behavior in the school area until 1958. This behavior involved running away, arriving late or completely missing class, very short attention span, poor motivation and distractibility. Richard progressed very slowly in all areas, attaining a reading level at one year six months and third grade arithmetic level by February 1958. Evidence suggested CNS involvement was indicated by abnormal electroencephalograph, poor motor control, low attention span, and by his gross inability to copy the Bender-Gestalt figures correctly.

When Richard entered the experimental group in February 1958, his Reading Vocabulary was 2.1; Reading Comprehension 1.0 and Total Reading Score 1.6. In June 1958, his Reading Vocabulary was 3.4, Reading Comprehension 2.5 and Total Reading Score 3.0, a gain of one year and four months. The following report was made by Richard's teacher[3] at that time: "Richard made more than substantial gains in all phases of reading and language. I attribute the greater portion of his success to the auditory reinforcement he received during the experimental period. Not only did Richard's grade level improve, but as he discovered he was able to grasp what he previously confused, his attention span and drive to learn grew stronger, thus enabling him to go forward."

Reading aloud, use of phonetics, block letters, tracing, color vowels, sand writing, and use of slit cards revealing only one line of reading material at a time, all seemed especially valuable. However, even with this special program, Richard was not doing well for a boy his age. He continued to perceive and transcribe units in reverse or upside down, but he seemed better able to recognize his errors, and spent more time constructively trying to remedy his difficulties. In general, his school behavior showed considerable improvement. He became more relaxed, developed a more positive attitude toward school work, and seemed to enjoy and benefit from the kinesthetic aids that had been incorporated into his school program.

Discussion

The children utilized in this investigation have a combination of difficulties so that they present no pure culture of any one group. They all have severe emotional problems, and this fact usually presupposes that they have had school problems. Although this study was not primarily concerned with the relationship between emotional difficulties and reading performance, the findings do indicate that perhaps any approach, which can be individualized and which stresses novel techniques so that the child is stimulated and motivated to learn, may be fruitful.

A question may be raised concerning the rather low IQ's obtained on the children. It has been recommended that in properly evaluating reading retardation, mental age be used and not chronological age. This was not done in the present study. The reasons for this are twofold. In order for children to be admitted to Bradley, the psychologist states whether the prospect is potentially of normal intelligence. Thus, not too much reliance is based upon the IQ alone as indicative of the child's intelligence. The fact that he is disturbed, plus the usual history of poor school performance, reduces the validity of the IQ. Intelligence was not by any means ignored as both groups were fairly closely equated for intelligence—at least, for responses to standard tests for "intelligence". It can be assumed that for a group of 12 the potential IQ should be about the same in both groups where the obtained IQ is approximately equal. At best any quantification of reading retardation is only a limited approximation. There are many factors which cannot readily be taken into consideration in setting up a quantitative evaluation of retardation. For example, many of the children missed one or two years of school because of their personality difficulties. Others may have been exposed to very little remedial work and if they had trouble in school were isolated and little time spent with them. Thus, it becomes almost impossible to make an accurate judgment of reading retardation. However, by setting up two fairly equated groups it can be supposed that any inequalities will be controlled by the groupings.

In considering the children who fall under the category of brain injured, it is very difficult to define clearly the nature and extent of the damage. The words "organic", "brain damage", or "cerebrally handicapped" have all been employed. The children at Bradley do not present severe neurological damage. A neurologist or a pediatrician often will not describe these children as being unusual. On occasion, however, they may be described as having awkward gait, or poor motor control. The EEG is not always useful in this respect because there may not be any cortical damage and usually the EEG findings may suggest only a delay in maturation. It is a most complex task to diagnose these "minimally brain damaged" children. They can best be screened by their behavior, school performance and visual motor response, and by a person with considerable clinical experience. The fact that many children were placed in the questionable category indicates that there is need for much further work in this area before increased diagnostic precision will be obtained. When there was some suggestion of brain damage but not enough to make a definite decision, the child was placed in the questionable category.

The major purpose of this report is to indicate a method that has been found most useful in teaching children who have minimal cerebral dysfunction. Much more work remains to be done. It may be that the procedures developed are most fruitful for emotionally disturbed children with cerebral dysfunction and do not apply to those of such children who are not emotionally maladjusted. This can be determined only by trying these procedures with such children who have no obvious maladjustment in a classroom setting.

The experimental aspect of this report is, of course, very tentative and the findings must be viewed as preliminary due to the small sample of the study. The investigation is being continued, but due to the small turnover in the Bradley population, it will not be possible to increase greatly the size of the sample in the immediate future. Therefore, it seems desirable to conduct a similar study in a regular public school setting where special classes can be set up for minimally brain damaged children and where different methods can be compared and evaluated.

In spite of certain methodological shortcomings in this research, there are two special features of the present study which may make it of some unique significance. One has to do with the fact that all the children in the study, experimentals and controls alike, were emotionally disturbed. They all were, and for a considerable period of time had been, receiving the benefit of all the resources of a children's psychiatric hospital—residential treatment center. This included individual psychotherapy, group approaches to treatment, medications where indicated and a host of other things.

This point is emphasized because the comment is often offered that apparent improvement from special educational procedures is actually a function of another process, of a psychotherapeutic nature, resulting from the new interpersonal relationship with the teacher. In the present study, however, the children were saturated with professionally organized and directed interpersonal relationships before the experiment began.

Another possibly significant aspect has to do with the suggestion that apparent results actually come from the close new relationship between teacher and pupil when programs of special education are introduced. Here, however, the children were already in very small, individualized classes, and a close intimate relationship with the teacher existed before as well as during the experiment. Moreover, the same teacher had both experimental and control classes. Thus, it seems that this exploratory investigation was based on an experimental design that is subject to less criticism and should lead to less equivocal results than those that have been obtained in previous studies in this area.

Summary

This paper reports on a somewhat unique approach to reading retardation involving kinesthetic and auditory aids applied to children with central nervous system involvement who also are emotionally disturbed. The preliminary findings contained in this pilot study appear to indicate the fruitfulness of this approach. The results reveal that children with CNS dysfunction taught by this method, for a three-month period, gained over a year in reading skills. This is in comparison with another such group of children who were taught by traditional methods. Their gain was four months. As the sample was rather small, the findings only suggest the feasibility of utilizing these methods for children with CNS dysfunction. A description of the method and techniques used was provided so that it would be possible for other teachers to practice the same methods in their own settings. A case illustration was given in order to indicate the application of the method and to emphasize the relationship between reading improvement and emotional adjustment.

Outline of Kinesthetic, Visual and Auditory Aids

1. Tracing words and letters
    a. Used mostly with trouble words, new words, and letters which the child repeatedly reversed
    b. All words were large and heavy, usually made by the adult with black crayon. Words were given an abundance of space; no more than three or four words on an 8 x 10 sheet of paper.
    c. A tracing paper overlay allowed the child to see clearly the words on the under sheet.
    d. At times dots replaced the tracing paper. A sample word would lead, then the game word lettered with dots would follow. By connecting the dots the child would complete the intended word.
2. Clay tray and sandbox writing
    a. Colored clay flattened evenly in a shallow tray and a stylus or pencil proved sufficient. A box 2 1/2 feet wide, 3 feet long and 4

inches deep filled with wet beach sand was desirable for sand writing.
   - b. Writing with a stylus or pencil in the clay offered much resistance and required time to replace the clay for the next exercise. These features tended to frustrate the children.
   - c. Writing in the sandbox with a wide stick, hand, or one finger produced more spontaneous letters and the children remained attentive for a longer span than they did while using the clay tray.

3. Air writing and blackboard writing
   - a. In both instances large sweeping strokes were advised. The children's letter form improved greatly when the sweep of their own arm was supposed to be the traveling of an airplane. An airplane noise was added which appeared to be supportive.
4. Block letters
   - a. Wooden block letters 4 inches high and 3/4-inch thick with a sand surface were used for word recognition, spelling and as a manipulatory device during a phonics lesson.
   - b. Each of the vowels were assigned a color while the consonants were the same. The child had a color cue for the vowels, was able to run his fingers along the rough side of the letter, and each letter had its own wooden mold from which it was cut, thereby having a positive block letter and a negative or depressed letter to fit.
5. Blockout cards
   - a. A card with a cut-out large enough to allow one printed line in the child's reader was used extensively for silent and oral reading.
6. Tachistoscope methods
   - a. Large flash cards and slide projected words were flashed before the children for approximately two seconds. The children would then give the word if they recognized it. Flash cards were also used for the basis of word games and for the children to read aloud as review.
   - b. Trouble words would be explained and returned to at the end of the game or session.
7. Phonics
   - a. Phonics was applied throughout. Emphasis was placed on letter sounds. Tachistoscope aids, block letters, tracing and sand writing were used in conjunction with phonics.
8. Other visual and auditory aids
   - a. Word meaning and sentence meaning were strengthened with auditory description and illustrative pictures. Blackboard drawings by the adult were most successful. This method proved to be non-time consuming and intriguing to the children.
   - b. The children were taught to spell aloud all trouble words when they were reading. By spelling the word aloud recognition was made easier.

FOOTNOTES

1. This study was made possible through the facilities and personnel available at the Bradley Hospital, where Dr. Talmadge is Senior Psychologist; Dr. Davids is Chief Psychologist and Dr. Laufer is Director. These authors wish to express their appreciation to the Bradley school teachers, Mrs. Loretta Foley, Mrs. Phyllis Murray, Miss Joan Barney, and Mr. Dwight Miller, without whose cooperation and assistance this study would not have been possible.
2. Detailed description of the pre-reading program and a lesson plan, containing specific examples of materials and procedures, have been mimeographed and will be sent to interested readers upon request.
3. The school report on Richard was furnished by Mr. Dwight Miller.

REFERENCES

1. Burks, H. F. The Effect on Learning of Brain Pathology, Exceptional Children, 1957, 24, pp. 169-172.
2. Durrell, D. D. Improvement of Basic Reading Abilities Yonkers: World Book, 1940.
3. Fernald, G. M. Remedial Techniques in Basic School Subjects New York: McGraw-Hill, 1943.
4. French, E. L. Kinesthetic Recognition in Retarded Readers, Educational Psychological Measurement, 1953, 13, pp. 636-654.
5. Laufer, M., Denhoff, E. and Solomons, G. Hyperkinetic Impulse Disorder in Children's Behavior Problems, Psychosomatic Medicine, 1957, 19, pp. 39-49.
6. Rabinovitch, A. L., Drew, R. N., DeJong, W. I., and Withey, L. A Research Approach to Reading Retardation, Neurology and Psychiatry in Childhood, Vol. XXXIV, Baltimore: Williams and Wilkins, 1956.
7. Strauss, A. A. and Kephart, N. C. Psychopathology and Education of the Brain-Injured Child, Vol. II, New York: Grune and Stratton, 1955.
8. Strauss, A. A. and Lehtinen, L. Psychopathology and Education of the Brain-Injured Child, New York: Grune and Stratton, 1947.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 6, February 1963)

A Comparison and Evaluation of Two Methods of Organization For the Teaching of Reading

REGINALD KIERSTEAD
Addison-Rutland District
Orwell, Vermont

IN ATTEMPTING to break the lock-step procedures which have characterized the graded elementary school in the past, two forms of organization for the teaching of reading have been employed in the Addison-Rutland District in Vermont.

Form 1: Ability grouping locked within the grade. Standards, materials, and methods are differentiated within the grade levels.

Form 2: All reading classes scheduled throughout the school at the same time. Students of like reading ability are grouped for reading instruction. Grade lines are entirely removed from the reading program. Standards, materials, and methods are adapted to the reading levels.

Question 1: Which of the two types of organization for the teaching of reading points to greater progress and improvement in vocabulary skills and reading comprehension skills?

Question 2: When groups are equated by means of scores received on the non-verbal section of the Pintner Test of General Ability, and by means of composite scores received on the Iowa Tests of Basic Skills, which of the following classifications of students shows most progress under the graded or ungraded structure?

1. The exceptional (130 IQ and above)
2. The above average (110-129 IQ)
3. The average (90-109 IQ)
4. The below average (70-89 IQ)
5. The exceptional (69 IQ and below)

Question 3: What effect does each type organization have on the ranges of vocabulary and reading comprehension skills as measured by the Iowa Tests of Basic Skills, Forms 1 and 2?

Procedure for Statistical Analysis

The following instruments and procedures were used to measure the results of the two programs:

1. Pintner General Ability Test (non-verbal) Grades 3-8.
2. Iowa Tests of Basic Skills (Form 1) Grades 3-8
3. Iowa Tests of Basic Skills (Form 2) Grades 3-8
4. Student Evaluation Sheets, Ungraded 7-8

Tables were constructed to show the percentages of students below and above the grade levels in September as compared to those below and above grade level in May. Tables were also constructed to show variations in ranges of the two skills in September as compared to May, the means for each group, the mean gains and losses, the percentages of students at the various intelligence quotient classifications, and the results of equating the groups by non-verbal intelligence quotients and September 15, 1959, composite scores from the Iowa Tests of Basic Skills.

Organizational Procedures, Ungraded Structure

Grades one and two were omitted from the plan due to the fact they are, and have been, ungraded. Because our reading system is based on the sequential development of skills, the basal readers were to serve as foundations for each group. Enrichment activities would be stressed.

All reading periods were scheduled at the same time. Students were assigned to reading levels on the basis of reading abilities, intelligence quotients, and teachers' judgments. Eleven reading level groups were established. During the reading period, each child moved to the room in which his reading level was being taught. Those students who were assigned to levels lower than their grades were assigned basal readers other than the readers used throughout the school so that the element of boredom would be eliminated.

Only students who were enrolled from September 15, 1959, to May 15, 1960, and who were tested at both times, were included in the study.

Results

When the groups were not equated, the data (see Table I) indicated that the students in the ungraded structure made more progress during the eight-month period in vocabulary skills and reading comprehension skills than students in the graded structure.

TABLE I

PERCENTAGES OF STUDENTS ABOVE GRADE LEVEL IN VOCABULARY SKILLS AND READING COMPREHENSION SKILLS AS MEASURED BY IOWA TESTS OF BASIC SKILLS ADMINISTERED MAY 15, 1960

Grade	% Graded Vocabulary Skills	% Ungraded Vocabulary Skills	% Graded Reading Comprehension	% Ungraded Reading Comprehension
3	74	87	57	56
4	54	45	49	48
5	65	75	62	70
6	49	71	38	47
7	42	50	53	55
8	45	44	45	37

TABLE II

MEAN GAINS OR LOSSES IN VOCABULARY SKILLS AND READING COMPREHENSION SKILLS AS MEASURED BY IOWA TESTS OF BASIC SKILLS ADMINISTERED SEPTEMBER 15, 1959 AND MAY 15, 1960*

Structure	Intelligence Quotients, Classes	N	Vocabulary Skills, Mean Gain or Loss	Reading Comprehension Skills, Mean Gain or Loss
Graded	Exceptional 130 and above	3	Gain 1 Year 1 Month	Gain 8 Months
Ungraded		3	Gain 1.6 Months	Gain 1 Year 3 Months
Graded	Above Average 110-119	51	Gain 1 Year	Gain 9.5 Months
Ungraded		44	Gain 9.2 Months	Gain 9.7 Months
Graded	Average 90-109	87	Gain 9 Months	Gain 8 Months
Ungraded		41	Gain 9.9 Months	Gain 7 Months
Graded	Below Average 70-84	13	Gain 7.5 Months	Gain 5 Months
Ungraded		14	Gain 6.4 Months	Gain 6 Months
Graded	Exceptional 69 and below	2	Gain 4 Months	Loss 3 Months
Ungraded		0		

*National Mean Gain: 8 months. Groups equated and classified by intelligence quotients as measured by the Pintner Test of General Ability, non-verbal.

TABLE III

MEAN GAINS OR LOSSES IN VOCABULARY SKILLS AND READING COMPREHENSION SKILLS AS MEASURED BY THE IOWA TESTS OF BASIC SKILLS ADMINISTERED SEPTEMBER 15, 1959, AND MAY 15, 1960*

Structure	Composite Score, September 15, 1959	N	Vocabulary Skills, Mean Gain or Loss	Reading Comprehension Skills, Mean Gain or Loss
Graded	1.9 to 1 year below Grade Level	12	Gain 8.9 Months	Gain 1 Year
Ungraded		4	Gain 9.7 Months	Gain 5.5 Months
Graded	.9 below Grade Level to Grade Level	55	Gain 8.8 Months	Gain 6.9 Months
Ungraded		32	Gain 6.5 Months	Gain 7.2 Months
Graded	Grade Level to .9 above Grade Level	63	Gain 8.6 Months	Gain 8.2 Months
Ungraded		40	Gain 9.6 Months	Gain 8.8 Months
Graded	1 year to 1.9 above Grade Level	20	Gain 9.5 Months	Gain 9.8 Months
Ungraded		24	Gain 8.3 Months	Gain 1 Year
Graded	2 years to 2.9 above Grade Level	15	Gain 7.9 Months	Gain 9.2 Months
Ungraded		9	Gain 8.3 Months	Gain 6.4 Months
Graded	3 years to 3.9 above Grade Level	1	Gain 6.0 Months	Loss 3.0 Months
Ungraded		2	Loss 1.5 Months	Gain .5 Months

*National Mean Gain: 8 months. Groups equated and classified by the composite scores received on the Iowa Tests of Basic Skills administered September 15, 1959

TABLE IV

THE RANGES FOR THE UNGRADED AND GRADED GROUPS AS MEASURED BY THE IOWA TESTS OF BASIC SKILLS ADMINISTERED ON SEPTEMBER 15, 1959, AND MAY 15, 1960

(Measured in Years and Months)

	Ranges: September 15, 1959				Ranges: May 15, 1960			
Grade	Graded Vocabulary	Ungraded Vocabulary	Graded Reading Comprehension	Ungraded Reading Comprehension	Graded Vocabulary	Ungraded Vocabulary	Graded Reading Comprehension	Ungraded Reading Comprehension
3	4.3	1.7	4.9	3.4	3.5	2.3	5.7	3.4
4	5.0	2.5	5.6	2.8	4.8	3.9	7.2	5.3
5	4.7	6.1	4.9	6.8	5.6	4.3	5.8	6.2
6	3.5	6.6	3.6	4.7	5.8	5.3	4.9	4.6
7	6.5	5.2	6.6	3.8	6.4	5.5	5.7	3.8
8	7.0	7.3	5.2	4.2	7.1	6.3	5.1	5.3

TABLE V

STUDENT EVALUATION OF ABILITY GROUPING

Questions	Better Group		Slower Group	
Have you felt able to keep up with the group? Yes? No?	Yes No	18 0	Yes No	19 1
Have you had the opportunity to take part in class discussions? Same as last year? More than last year?	Same as More than	9 9	Same as More than	6 11
How would you rate yourself within the group? Upper third? Middle third? Lower?	Upper third Middle lower	7 10 1	Upper third Middle Lower	5 13 1
How do you think your progress compares with last year? Better? The same? Poorer?	Better The same Poorer	11 5 2	Better The same Poorer	16 3 0
If you could select your group next year, would you choose the same or would you prefer to be in the other?	The same The other	17 1	The same The other	5 10

When the groups were equated by intelligence quotients (Table II), there were no significant differences between the gains made in vocabulary skills and reading comprehension skills by the two groups.

When the groups were equated by the composite scores of the Iowa Tests of Basic Skills administered on September 15, 1959 (Table III), there were no significant differences between the gains made in vocabulary skills and reading comprehension skills by the two groups.

Student Evaluation of Ability Groupings

Students in the seventh and eighth grades who were grouped according to ability were asked to answer the five questions listed in Table V. No signatures were required on the questionnaires. They were completed by the students during one class period; therefore, there was no opportunity for the students to discuss the questions with each other.

The students' evaluations of ability groupings indicate that they feel comfortable academically as a result of being separated by abilities; however, while the faster group does appear to feel comfortable socially, the slower group does not. Two-thirds of them would transfer.

Conclusions

1. Classifying students by ability cannot in itself remove individual differences or the need for adapting instruction to individual differences.

2. Due to the narrower ranges of abilities in classrooms organized by ungraded plans, materials can more easily be assembled and standards and methods more easily adapted to meet individual differences.

3. Whether the graded structure with ability grouping or the ungraded structure is used, there must be constant evaluation and study of the individuals comprising the groups, and groups must be kept flexible and elastic.

4. Parents and students accept ability groupings if lines of communications between homes and schools are kept open.

5. Teachers' judgments are as important criteria for grouping as are test results.

6. Teachers prefer the ungraded structure.

Table IV shows that although range is an unstable measure of dispersion, it was used to determine if scores in the low extremes and the high extremes were greatly affected by either program. Such does not appear to be the case. The mean increase of ranges for the graded group in vocabulary skills was four months of achievement, and for the ungraded, one-half month. In reading comprehension skills for the graded group, the mean increase of ranges was six months and for the ungraded, five months. Although the Table shows no significant changes in ranges confronting teachers in heterogeneously grouped classes, and the necessity for some form of ability grouping in order to narrow ranges. The table also shows that the higher the grade, the greater the ranges.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 6, February 1963)

The Relationship Between Intelligence Quotients and Gain in Reading Achievement with Arithmetic Reasoning, Social Studies, and Science

CARRIE M. SCOTT
Bend, Oregon

A NUMBER OF investigations have been made in the field of education, correlating intelligence with proficiency in the various academic fields. Other investigators, including Artley (1), Coffing (2), Fay (3), and Hinkleman (5) have been interested in reading as it relates to other curriculum areas. All have found that a positive relationship exists. A few studies have been made to ascertain whether there is an existing relationship between intelligence and gain in achievement in the curriculum areas, or whether a gain in reading achievement is concomitant with a gain in achievement in the other areas. Tilton (9) computed gains within a group of fourth and fifth-grade pupils over a 20-month period making intercorrelations between the sub-tests of the Unit Scales of Attainment. Teachers are inclined to believe that when a pupil makes a gain in reading achievement, he will gain in other subject areas to the same extent.

The Problem

The purposes of this study were to learn 1) whether a relationship exists between intelligence scores and gain in reading, arithmetic reasoning, social studies, and science achievement, and 2) whether a gain in reading achievement is related to a gain in arithmetic reasoning, social studies, and science achievement.

Procedure

To accomplish these ends, three groups totaling 670 pupils in the Bend Public Schools for whom intelligence test scores were available, and who had taken the Stanford Achievement Test, Intermediate Battery (6) at the beginning of grade five and again at the end of grade six, were selected for study. The test data were obtained as follows:

Fall, 1956, Form N ----- Spring, 1958, Form K
Fall, 1957, Form J ----- Spring, 1959, Form N
Fall, 1958, Form M----- Spring, 1960, Form L

The intelligence scores were derived from the California Test of Mental Maturity (7). The series used depended upon the grade level in which the tests were administered. If a Stanford-Binet Intelligence Scale (8) or a WISC (10) had been administered this score was preferred.

For each pupil a record was made of the intelligence quotient, the grade score in reading, both paragraph and word meaning, arithmetic reasoning, social studies, and science. Referring to the table in the testing manual (6) the grade scores were converted into K-scores. The K-scores were averaged for the paragraph and word meaning tests to find an average reading score. Differences in K-scores were found for each pupil over the 19-month period.

The grade scores on the tests were converted to K-scores to obtain more accurate measurement. The K-scores have units which are approximately equal throughout the range of the tests and, therefore, reveal growth patterns with greater accuracy than those obtained through the use of grade equivalents. The differences in performances on the two batteries more nearly represent true differences when K-scores form the bases for making the computations (6: Manual).

Over the three-year period the fifth grades had twelve teachers; six having taught in the fifth grade for three years, two for a period of two years, and four for one year. One of the fifth-grade teachers went into the sixth grade for the last two years of the study. Four of the sixth-grade teachers had taught in the grade for three years, three for four years, three for two years, and five for one year, making a total of twenty-three teachers for the two grades.

Grade scores were averaged to compare the achievement scores of the pupils in the two grades with the national norms.

Grade 5,	
National Norm, grade 5. 1	
Average composite reading grade	5. 7
Average arithmetic reasoning grade	5. 1
Average social studies grade	5. 5
Average science grade	5. 7
Grade 6,	
National Norm, grade 6. 8	
Average composite reading grade	7. 3
Average arithmetic reasoning grade	7. 3
Average social studies grade	7. 4
Average science grade	8. 1

According to the average grade levels obtained on the Stanford Achievement Tests, the pupils in these groups were able to perform at slightly above the expected grade levels.

The mean intelligence quotients and the mean gains in the four curriculum areas are shown together with their standard deviations in Table I.

The mean intelligence quotients indicated that the group rated at slightly above the average capacity. The highest gain was in arithmetic reasoning, followed by gain in social studies and science, with reading gain being the lowest.

Using the Pearson Product-Moment formula (4) correlations were found between intelligence scores and gain in reading, arithmetic reasoning, social studies, and science achievement. The results of these investigations appear in Table II.

All of the obtained correlations were low, but positive, except for the small negative correlation between intelligence scores and gain in science. The correlation between gain in reading and gain in science also was very low.

In order to determine the relationship existing between gain in reading achievement and gain in the three other curriculum areas, partial correlations were computed with the effect of the intelligence quotients held constant (4).

Gain in reading achievement correlated with-

gain in arithmetic reasoning	.153
gain in social studies	.159
gain in science	.086

When the intelligence quotients were held constant and did not interfere, the correlations were positive, although low.

It was observed that some pupils neared the top of the achievement tests at the beginning of grade five and did not advance much by the end of grade six. Therefore, an effort was made to determine whether the test covered the materials being taught and whether it was long enough and/or difficult enough for the pupils at the end of the sixth grade. To this end the brief questionnaire which follows, concerning the areas of reading, arithmetic reasoning, social studies and science, was checked by the sixth-grade teachers.

1. How well do you feel that the Stanford Achievement Test administered at the end of grade six covers the materials presented in your textbooks during the year?
   A. Well _____; B. Fairly well _______;
   C. Poorly______
2. Is this test long enough or difficult enough so that-
   A. No one finishes within the time limit?___
   B. A few pupils complete it within the time limit? ____
   C. The majority have completed it within the time limit? ______
3. Achievement-
   A. No one receives a perfect or nearly perfect score.______
   B. A few pupils receive perfect or nearly perfect scores.______
   C. A number of pupils receive very high scores.______

Concerning the reading test, the teachers reported that the majority of the pupils had finished within the time limit, and that many received very high scores.

In arithmetic reasoning, they felt that the test covered the materials taught during the year "well" or "quite well", and that several pupils finished within the time limit. They varied in reports of achievement from "no one receives a very high score" to "several pupils receive very high scores".

The social studies material was reported as being poorly adapted to the content of the sixth-grade course of study, but the majority of pupils had finished the test within the time limit, and many had received very high scores. The teachers were not in agreement as to how adequately the science items were treated in the instructional procedures. They agreed that the majority had finished the test within the time limit and had made very high scores. One teacher wrote that an achievement test of this nature would of necessity be general in content and not expected to cover materials taught in a specific grade.

With the exception of the social studies test, the sub-tests studied in the battery seemed to cover the materials taught by the end of the sixth grade. Since many of the pupils completed the tests ahead of the time limit and made high scores it may be concluded that the sub-tests did not have a high enough "top" for the brighter pupils. A more accurate measurement of those who made high ratings at the beginning of the fifth grade may be obtainable by administering the more advanced battery. At its best an achievement test can be only a sampling of what a pupil knows, and a number of factors may influence the score a child makes on an individual test.

It was believed that if the pupils were divided into groups according to intelligence quotients, the relationships between gain in reading achievement and gain in the other curriculum areas might differ among those of different mental capacities. Consequently, 390 of the subjects were used whose intelligence quotients fell into one of three groups: the low capacity group, 70-85; the average capacity group, 95-110; and the high capacity group, 120-135. By using only the pupils whose intelligence ratings fell into one of the three groups the mean intelligence quotients and mean gains in achievement were found with their standard deviations, and are shown in Table III.

The highest capacity group made the greatest amount of gain, but was surpassed by the average group in gain in science. The lowest capacity group made the least amount of gain although the high group was surpassed with respect to gain in science.

Correlations were obtained between intelligence scores and gain in reading achievement, and between reading achievement and gain in arithmetic reasoning, social studies and science achievement. The results of the correlations appear in Table IV.

The correlation between reading achievement and social studies achievement was high with the low capacity group. Otherwise, the correlations were low, but positive in every case. The correlation between gain in reading and gain in science was lowest with the average group, and the correlations between intelligence quotients and gain in reading and gain in social studies were lowest with the high capacity group.

Discussion

Research studies have shown that there is a correlation of .50 or higher between intelligence test scores and scores on subject matter tests. When intelligence scores are correlated with gain in academic achievement the correlations are lowered considerably. Low correlations are to be expected between gains since gains are less reliable than the scores used to compute them.

Fay (3) believed that the reading skills significant for success in one subject may differ for success in another, and to study general reading ability as it relates to other curriculum areas is inadequate. Artley (1) felt that reading skills are both general and specific, but that the social studies are most adapted to general reading ability. The obtained correlation of .80 between gain in reading and gain in social studies with the low capacity group confirms this premise. Coffing (2) found positive correlations between reading ability and improvement in arithmetic. Hinkleman (5) believed that reading had an important relationship to elementary school achievement, but he felt that intelligence also must be considered. Tilton (9), in making intercorrelations between measures of school learning, found the correlations between reading and arithmetic fundamentals, social studies, and science to range from .217 to .323. He found that reading gains were less closely associated with progress during the fourth and fifth grades than in the lower grades, and this was due perhaps to the more able pupils acquiring the skill earlier. Therefore, bright pupils did not continue to improve in ability at the earlier rate of progress. He believed that progress in the information areas during the fourth and fifth grades resulted from functioning of the ability to read already present at the beginning of the period. Since the present study concerns itself with the pupils in grades five and six, his conclusions may apply. Tilton stated that any achievement test was an inadequate measure of the results of good teaching, or of knowledge held by a capable pupil. Since general achievement tests are short and general in content, this criticism is believed justifiable.

It is a known fact that intelligence has a causal relationship to school learning, and also that one must learn to read in order to succeed with informational materials. Since the relationships were present when the intelligence scores were held constant, both intelligence and reading proficiency were influential factors in academic progress. Also, since there still were positive correlations when the intelligence scores were held constant, we know that intelligence was not the only factor involved in the relationships. It can be assumed that reading plays a part in academic achievement as well as other factors.

The pupils involved in the study had a total of twenty-three teachers all of whom were well-trained, efficient, and for the most part experienced. With any group of teachers there is a variation in teaching ability in instructing in different subject areas. A sufficient number of teachers and pupils were involved to form heterogeneous groups, and thus, the variations were lessened.

Findings

1. There were wide variations in the amount of gain as measured by the achievement tests for individuals of similar intellectual status.
2. Individual pupils were inconsistent as to the amount of gain among the different sub-tests of the battery.
3. Average grade-scores on the achievement tests indicated that the pupils rated slightly above the national norms.
4. When 670 pupils were used, the highest correlation was between intelligence and arithmetic reasoning, and the lowest correlation was between intelligence scores and science.
5. Gain in reading achievement correlated highest with gain in social studies and lowest with gain in science.
6. Many pupils of high capacity rated high on the achievement tests at the beginning of the fifth grade and did not make a large gain at the end of the sixth grade. A test with a higher "top" is indicated for the spring testing in grade six.
7. The correlations were lowered when the intelligence test scores were not complicating the investigation.
8. There is a positive correlation not attributable to intelligence between reading gain and gain in arithmetic reasoning, social studies, and science.
9. These data are consistent with the belief that improved reading ability may lead to improvement in arithmetic reasoning, social studies, and science, but the degree of concomitance is less than many educators realize.
10. The teaching of reading needs to be general in developing skills and affording practice, and must also be specifically related to the other curriculum areas.

REFERENCES

1. Artley, A. Sterl. "A Study of Certain Relationships Existing Between General Reading Comprehension and Reading Comprehension

in a Specific Subject Matter Area," Journal of Educational Research, XXXVII (February 1944), pp. 464-473.

2. Coffing, Esther A. "The Relationship Between Silent Reading Ability and Arithmetic Ability," School Science and Mathematics, XLI, (January 1941).
3. Fay, Leo C. "The Relationship Between Specific Reading Skills and Selected Areas of Sixth Grade Achievement," Journal of Educational Research, XLIII (March 1950) pp. 541-547.
4. Garrett, Henry E. Statistics in Psychology and Education (New York: Longmans, Green and Co. Inc., 1958).
5. Hinkleman, Emmet A. "Relationship of Reading Ability to Elementary School Achievement," Educational Administration and Supervision, XLII (February 1956), pp. 65-67.
6. Kelley, Truman L. and others. Stanford Achievement Test (Chicago: World Book Company, 1953).
7. Sullivan, E. T., Clark, W. W., and Tiegs, E. W. California Test of Mental Maturity (Los Angeles: California Test Bureau, 1953).
8. Terman, Lewis M. and Merrill, Maude A. The Revised Stanford-Binet Scale (Chicago: Houghton Mifflin Company, 1937).
9. Tilton, James Warren. "The Intercorrelations Between Measures of School Learning," The Journal of Psychology, LXXXV (April 1953), pp. 169-179.
10. Wechsler, David. Wechsler Intelligence Scale for Children (New York: The Psychological Corporation).

TABLE I

MEANS WITH THEIR STANDARD DEVIATIONS ACCORDING TO INTELLIGENCE QUOTIENTS AND GAINS IN ACHIEVEMENT IN FOUR CURRICULUM AREAS ACCORDING TO K-SCORES FOR 670 PUPILS OVER A 19-MONTH PERIOD

Areas	Means	Standard Deviations
Intelligence Quotients	108.30	13.55
Reading Achievement	8.35	5.13
Arithmetic Reasoning	13.70	7.05
Social Studies	11.34	7.05
Science	9.96	6.51

TABLE II

CORRELATIONS BETWEEN INTELLIGENCE QUOTIENTS AND GAIN IN READING, ARITHMETIC REASONING, SOCIAL STUDIES AND SCIENCE ACHIEVEMENT, AND BETWEEN GAINS IN READING AND ARITHMETIC REASONING, SOCIAL STUDIES, AND SCIENCE TOGETHER WITH THEIR PROBABLE ERRORS

Areas	Correlation	Probable Errors (Plus or Minus)
Intelligence and		
Gain in Reasoning	.239	.025
Gain in Arithmetic Reasoning	.358	.024
Gain in Social Studies	.338	.023
Gain in Science	-.031	.026
Gain in Reading and		
Gain in Arithmetic Reasoning	.224	.025
Gain in Social Studies	.231	.025
Gain in Science	.076	.025

TABLE III

THE MEAN INTELLIGENCE QUOTIENTS AND MEAN GAINS IN ACHIEVEMENT IN FOUR CURRICULUM AREAS ACCORDING TO K-SCORES TOGETHER WITH THEIR STANDARD DEVIATIONS FOR 390 PUPILS GROUPED ACCORDING TO INTELLIGENCE QUOTIENTS

Areas	Intelligence Quotients					
	70-85		95-110		120-135	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
Intelligence Scores	80.97	7.96	103.63	8.62	126.60	7.98
Reading Gain	5.20	4.02	8.97	4.52	10.43	5.38
Arithmetic Reasoning	9.80	5.00	11.74	4.46	16.24	7.92
Social Studies	5.27	4.82	11.22	6.70	15.02	7.72
Science	10.70	5.36	11.26	6.32	9.96	6.96
Number of Pupils	30		269		91	

TABLE IV

CORRELATIONS BETWEEN INTELLIGENCE QUOTIENTS AND GAIN IN READING ACHIEVEMENT, AND CORRELATIONS BETWEEN GAIN IN READING ACHIEVEMENT AND GAIN IN ARITHMETIC REASONING, SOCIAL STUDIES, AND SCIENCE ACHIEVEMENT TOGETHER WITH THEIR PROBABLE ERRORS FOR 390 PUPILS DIVIDED INTO THREE GROUPS ACCORDING TO INTELLIGENCE QUOTIENTS

Areas	Intelligence Quotients					
	70-85		95-110		120-135	
	Correlation	P.E.	Correlation	P.E.	Correlation	P.E.
Intelligence Quotients and Reading	.324	.110	.195	.039	.010	.071
Reading and Arithmetic Reasoning	.284	.112	.314	.037	.163	.069
Reading and Social Studies	.803	.044	.231	.038	.079	.070
Reading and Sciences	.274	.113	.015	.041	.133	.069

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 6, February 1963)

The Effectiveness of Workbooks in the Elementary School Reading Program

ROBERT L. DOCTER
San Fernando Valley State College
Northridge, California

WIDELY divergent viewpoints exist concerning the value of reading workbooks in the instructional process of the elementary classroom. Opponents of the use of workbooks in the reading program criticize them as busywork and claim it is impossible to meet the individual needs of children with this type of material. Proponents identify the workbooks as a near panacea for the ills of the entire reading program and claim that there are vast savings in teacher preparation time when the workbooks are used. Strangely, in educational literature there seems to be little middle ground between the two factions and even less research evidence to support either viewpoint.

To determine what objective evidence could be obtained regarding the advantages or disadvantages of workbook or non-workbook methods, a study was undertaken by this writer in 1959. With the approval of the Los Angeles Board of Education, the objectives of this study were implemented by staff members of the Division of Intructional Services and Elementary Education.* This research, therefore, attempted to measure the relative effectiveness of workbook and non-workbook methods of follow-up activity through an analysis of empirical evidence gathered in a highly controlled learning experiment.

The "workbook method" was defined as a reading follow-up activity involving the use of a commercially published, consumable book designed to accompany a specific basal reader. A "non-workbook method" was defined as a reading follow-up activity involving the use of materials designed and prepared by individual teachers for specific reading groups. The term "follow-up," when used to describe a part of a lesson or type of reading material, was defined as a reading activity designed to reinforce previous teaching and to afford practice for the pupil in specific reading skills.

Procedure

Eighteen Los Angeles elementary schools were randomly selected from the population of approximately 450 elementary schools of the area for the purpose of providing a sample of intact classrooms. Each of the eighteen schools was randomly assigned one of the six elementary grades, and two classrooms of that grade level were included in the study. For example, school "A" furnished two first grades, school "B" furnished two second grades, and so on until all of the elementary grades in all of the schools were assigned. One of the assigned classes in each school was identified as an experimental group and utilized the workbook method of instruction for the entire school year of 1958-59. The other class in each school was identified as a control group and proceeded with the non-workbook method of instruction for the same period.

Through assignment of both experimental and control groups at each of the schools, it was possible to assume that the pupils in each classroom had similar socio-economic backgrounds, similar access to supervision and equal opportunity to use the library. Administrators of each school were asked to assure that a similar amount of time was devoted to the reading program and that class sizes in both the experimental and control groups were similar. Workbooks were selected to accompany all titles of two basic reading series in widespread use throughout the area as well as five other titles from three publishing houses whose books were used at specific grades.

Instruction proceeded in both the experimental and control group classrooms in accordance with established developmental methods practiced in the Los Angeles area. The only exception involved the use of the workbook in the experimental situations. Each teacher divided his class into three separate reading groups, based on the reading achievement scores from standardized tests, and assigned a basal reader to each group which corresponded with the specific instructional reading level involved. Growth in reading during the one-year experimental period was measured through use of the Standard Reading Achievement

---

* The administrative officer for the school district was Mrs. LaVon H. Whitehouse, Coordinator of Elementary Curriculum.

TABLE I

MEAN GAINS IN COMPREHENSION AND VOCABULARY BY GRADE LEVEL

Reading Skill	Group	Grade Level					
		1	2	3	4	5	6
Comprehension	Experimental	5.52	18.33	14.07	8.55	7.58	7.91
	Control	9.42	13.75	8.06	5.59	8.76	5.67
Vocabulary	Experimental	4.87	14.50	12.04	7.95	8.27	6.43
	Control	7.10	8.59	6.67	5.96	9.47	5.52

TABLE II

ANALYSIS OF COVARIANCE F RATIOS OF BETWEEN AND WITHIN SUB-GROUP MEAN SQUARES TO DETERMINE THE SIGNIFICANCE OF DIFFERENCES IN COMPREHENSION AND VOCABULARY GAINS BY GRADE LEVEL

Reading Skill	Grade Level					
	1	2	3	4	5	6
Comprehension	28.50**	6.33*	5.73*	5.16	.157	3.52
Vocabulary	7.55**	16.78**	8.68	3.17	.202	1.01
df	1,200	1,100	1,100	1.80	1,175	1,200
$F_{.01}$	6.76	6.90	6.90	6.96	6.77	6.76
$F_{.05}$	3/89	3/94	3.94	3.96	3.90	3. 9

** Significant beyond the .01 level of confidence.
 * Significant beyond the .05 level of confidence.

Tests. Form "J" was administered in the fall as the pre-test instrument, and Form "K" was administered in the spring as the criterion measure. The mental age of each pupil, derived from group intelligence measures, was also recorded as an additional antecedent variable.

Each classroom group was left intact and analysis of covariance was used to test the hypothesis that both experimental and control groups were random samples of the same population. This device is similar to that suggested by Walker and Lev (5) and Caffrey (2). Each grade level was treated as a separate experiment. A "gain" approach, also suggested by Caffrey (3), was utilized in the statistical design. In this instance, where one of the antecedent variables is a pre-test measure and the criterion variable is a post-test measure, the difference between the two measures becomes a record of the gain. Thus, the mental age of each pupil became the only antecedent variable and the statistical computation was simplified.

Findings

Mean gains during the school year of the experimental and control groups in raw score comprehension and vocabulary points are summarized in Table I. Gains in the first grade favored the control (non-workbook) group. With use of the mental age as the antecedent variable, the significance of this difference was tested through analysis of covariance. Table II presents the results of the F tests performed to determine this significance. An analysis of the first grade F ratios in Table II shows that the difference in gains between the experimental and control groups was highly significant beyond the .01 level of confidence in both comprehension and vocabulary scores.

In the second, third, and fourth grade columns of Table I, however, it is noted that the differences in gain favor the experimental (workbook) group. The second grade F ratios, as recorded in Table II, show that the comprehension gain of the workbook group was significantly greater than that of the non-workbook group beyond the .05 level, and that the vocabulary difference favored the workbook group beyond the .01 level. An identical situation was noted with regard to the third grade scores. There were significant differences in comprehension gains and highly significant differences in vocabulary gains, both favoring the experimental group. The fourth grade column of Table I also shows that the experimental group registered greater gains during the year than did the control group. According to the findings presented in Table II, this difference was significant in regard to comprehension, but not significant in terms of vocabulary. No significant differences were apparent at the fifth or sixth grade levels.

Interpretation

The significantly greater gains of the first grade control group suggest the need for a close analysis of the beginning reading program. On the basis of this evidence, it would appear that children in the early stages of reading benefit more from the intimate, closely related follow-up material prepared by the individual teacher for a unique classroom presentation than do children whose beginning reading experiences involve the use of readiness and pre-primer workbooks. The immediate and dramatic reversal of the results at the second grade demonstrates a vital need for the introduction of these materials at this point in the curriculum. Certainly the increased readiness of the child to learn independently in a formal way is one factor pointing to the significantly greater gains of the workbook group at this grade level. The highly significant differences in vocabulary scores at the second and third grades might possibly stem from superior practice material in word analysis skills found in the workbook exercises.

Summary

The purpose of this investigation was to measure the relative effectiveness of workbook and non-workbook methods of reading follow-up activities in each of the six elementary grades. With the mental age of the student as the antecedent variable, the gain between the reading pre-test and post-test score was used as the criterion variable. The hypothesis that both groups were random samples of the same population was tested through analysis of covariance. Statistically significant differences existed between the gains of workbook and non-workbook classes in both reading comprehension and vocabulary at specific grade levels. This difference favored the non-workbook classes at the first grade level and the workbook classes at grades two, three, and four. No significant differences were apparent at grades five and six.

REFERENCES

1. Betts, E. A., Foundations of Reading Instruction, New York, American Book Company 1954.
2. Caffrey, J., "Estimated True Growth: Lord's Equations Applied to Reading Test Data," California Journal of Educational Research VII, 1956.
3. Caffrey, J., Estimation of Learning as a Function of Measured Ability. (Resume of material presented at the annual conference of the California Educational Research Association, Sacramento, 1959.)
4. Russell, D. H., Children Learn to Read, Boston, Ginn and Company, 1949.
5. Walker, H. and Lev, J., Statistical Inference, New York, Henry Holt, Inc., 1953.

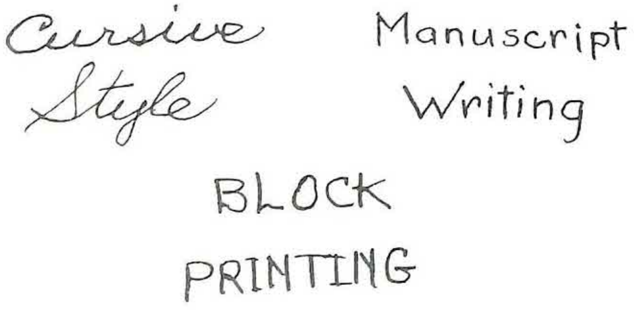
THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 6, February 1963)

Simplified Handwriting for Today

GERTRUDE HILDRETH
Brooklyn College

IN SPITE OF typewriters and telephones, handwriting continues to serve for convenient communication and record making. The older generation learned to write a flowing cursive hand by dint of persistent practice at school. From the early 1920's when manuscript style writing was introduced (a style similar to typing), school beginners learned these simple letter forms which could be used for expressive writing almost from the first lesson. Originally, the new style was presumed to be the pupil's permanent hand, with no suggestion that at midpoint in learning a changeover to cursive style writing would be required. However, as the new style passed the experimental stage and became widely adopted by public school systems, a transition to cursive style writing was required in the later primary period, sometimes as early as the end of second grade.

In the meanwhile, another change in writing usage was taking place. The business world with increasing frequency requested people to "print" important information in filling out application blanks and other documents.



Was the frequency of this request due to the greater volume of forms to be processed, or was it because we were becoming more careless in handwriting? Whatever the cause, this matter brings up the question of handwriting skills suited to our practical writing needs today, and the methods of school instruction required to produce the best results.

How do people respond when requested to "print" while filling out business papers? To the older generation, "printing" can mean only slow, painstaking block lettering, but younger people are inclined to interpret "print" as manuscript style writing learned in the primary grades, often referred to there as "printing." Some people simply turn to a handy typewriter when requested to print.

During recent years, clerks employed in stores in the larger cities have been taught block printing and are required to use it in making out all sales slips and other memoranda, a slow process for many, judging from the effort neat block printing requires.

In order to check up more fully on this situation, a study was made of the handwriting of college students who were requested to print while filling out registration cards for the fall term, 1961, at Brooklyn College, New York. This required the printing of full name, address, and other information. Records were obtained for two classes of junior year undergraduate students in education courses and two classes of graduate students in education courses, who, for the most part, had several years of teaching experience and training in elementary school work. Of the total of 103 students in these four classes, ten were men.

The table at the top of the next page indicates the handwriting response made by these students when they were requested to print. The cursive style writers must have forgotten the request to print. The junior year students were in the age range 18-19, the graduate students in their 20's and 30's, for the most part. The men uniformly block-printed except for one who typed. An inspection of the registration blanks proved that some of the letters in block printing were scarcely legible and the students had to be asked to give their names.

These results indicate that the undergraduates did block printing or wrote in manuscript style in about 50-50 proportions, while the graduate students showed more inclination to do manuscript style writing. The difference is due to the fact that graduate students in elementary education have brushed up on manuscript style or learned it in their methods courses in the senior year, and those who teach in the primary grades use it daily in classroom instruction, whereas the undergraduates have not had elementary methods courses and have

Class		Block Printing	Manuscript Style	Manuscript and Block Lettering	Cursive Style	Typing	Totals
Junior Year	I	16	14	1	1		32
Junior Year	II	16	12	1		1	30
Graduate	I	5	20	4	1		30
Graduate	II	3	26		2		31

had no experience in teaching handwriting. The men in these classes were high-school teachers lacking elementary methods and teaching experience in the elementary grades.

When the students were queried about handwriting instruction in the grades, most of these young people indicated that they had first learned manuscript style (printing) and had changed over before the end of the primary period.

A new bulletin, *Teaching Handwriting*, states that the changeover from manuscript style to cursive writing is standard practice throughout New York city except in special cases.[1]* For a few children, the physically immature, the left-handed, the slow learners, those with motor handicaps, the partially seeing children, the authorities recommend continuing with manuscript writing. Most children make the change in the third grade. Some children, this publication states, may not be ready until the fourth grade or later. The bulletin goes on to say that when the children have acquired facility in both styles, they use cursive for most writing activities and manuscript for those situations in which it is most appropriate. Block printing or lettering is widely taught in high school shop and drafting courses. In some cases, children have been changed back and forth more than once in moving from one city to another and even in elementary schools within the same school system.

Although the New York city bulletin properly uses the term "manuscript writing" in referring to primary instruction, teachers, parents, and the children themselves refer to this unjoined letter style as "printing", not only in New York city but throughout the country. The children are said to be learning "to print"; the term "handwriting" is not used. We are told, "Now they're only printing. Next year the children will be taught to write". This misleading terminology is unfortunate, since it gives the impression that manuscript style is not actual handwriting, and that it must inevitably be discarded.

Most people who elect to "block print" in responding to the request, "please print", write in an unaccustomed, unpracticed hand since block printing is not taught in the elementary school and only a few learn it in high school. Consequently, the writer must necessarily slow down and print laboriously.

Reasons for the Change-Over

Why is a change-over in style of handwriting universally required in the primary grades today? The reasons given are that manuscript style is slower than cursive style at maturity and that most parents object to the continuance of babyish printing, in part because it is said to be slower and also because it does not look "educated" enough for the middle and upper students. Another reason is that teachers in the intermediate and upper grades have had no instruction in teaching manuscript writing at an advanced level.

Research studies dating back 30 and 40 years, comparing the results of manuscript writing and cursive style, have demonstrated that manuscript style is not necessarily slower than cursive style with equal amounts of practice through the elementary grades.[2] There is no reason why manuscript should be slower if children beyond the primary grades practice for fluency and rhythm. The writer found that upper graders who had always used manuscript style and did extensive expressive writing actually wrote faster than their classmates who had always used cursive style.

The belief that manuscript style must be slower is due to the fact that children who learn and use manuscript have generally been "contaminated" at some time or other with cursive style instruction, and they have not been drilled as thoroughly to develop fluency as those who are taught the traditional cursive style.

The parents' request for the changeover is insistent only in the brief span of grades four to six or slightly over. They willingly accept manuscript writing for their children in the primary grades because of the demonstrated advantages for young children in the beginning stages of learning to read and write. They freely consent to a "change-back" to manuscript writing or to instruction in block lettering or typing in the junior high school to insure neater, more legibly written papers. Typical parents regard cursive style writing as a symbol of traditional academic schooling, the acquisition of a skill that demonstrates the child's increasing scholastic maturity. Manuscript style writing is regarded as an infantile stage which must be abandoned for a more grown-up accomplishment.

---

*Footnotes will be found at end of article.

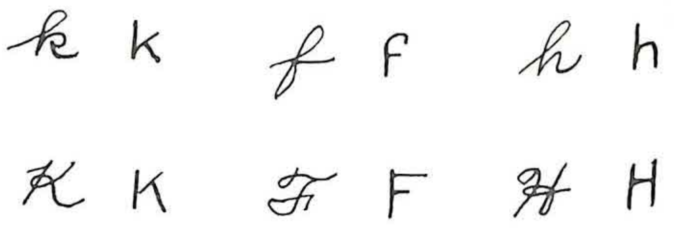
Another argument for the change-over is that teachers in the intermediate and upper grades have not been trained to build mature writing habits from the children's manuscript learned in the primary grades; hence, they are unable to assist the children in developing rhythm and fluency. Furthermore, there are few good teaching materials for manuscript style in the upper grades and no training courses that give instruction in teaching advanced writing in simplified styles.

In the days when the upper grade teacher's chief task in handwriting instruction was to help a child improve cursive handwriting habits that were fairly well established by the intermediate grades, the problem was to assist individual children who presented some special difficulty for one reason or another; but today, upper grade teachers must not only assist individual children with handwriting problems, but also teach a new style to the entire class.

Objections to the Change-Over

A change-over from manuscript style writing learned in the primary grades to cursive style beyond that point is both wasteful and unnecessary. Since the psychology of learning proves that first-learned skills tend to persist and to interfere with later-learned related skills, pupils must be thoroughly retaught and drilled in the new forms to prevent reversion to earlier forms and to offset interference of the earlier manuscript writing practice. The partially learned, hastily written cursive writing may break down or never really be perfected in the upper grades. The data cited above suggest how readily young people in teacher training courses can change back to manuscript style when occasion requires.

A massive reteaching effort is required to "retool" primary grade handwriting just at the time that children have gained considerable control over simple letter forms and are beginning to enjoy expressing their ideas in writing. After nearly three years or more of practice in simple, clear-cut letter forms they are required to learn a new style with different letter strokes and joined letters. There is no such thing as a natural transition from manuscript style writing to cursive, because the letter strokes in the two styles are distinctly different for at least half of the letters. Here are some illustrations:



Teachers say that the children are changed-over when they are "ready to make the change". However, no school-age child is ever as ready to change abruptly to a new motor habit as he is to continue in the accustomed motor patterns. A change-over at this point entails psychological interference. Although most children are successfully retrained within a short time, the children show by their behavior and comments that they resist the change, whether consciously or not.

Although at the point of the change-over the pupils have already had several years' experience in learning to write, now they must be treated more like beginners learning a new skill. This entails different problems for the upper primary and intermediate grade teacher than teaching the six-year-old beginner to write cursive style. Practice from the point of the change-over must be as thorough as in the early grades. Time that is needed in the middle and upper grades for improvement in expressive writing and written work in the content studies must be given to drill to perfect the new writing habits.

The Need of Simplified All-Purpose Handwriting Styles

The confused situation in present-day handwriting practices raises the question whether it is possible to devise a clean-cut style of handwriting that is highly legible, can be written fluently without fatigue; handwriting that need not be changed over in an intermediary stage, but can be developed to a point of high proficiency through systematic instruction by the end of the elementary school period. The answer is decidedly "yes", judging from observations of adults who learned manuscript style in the first place, practiced to perfect it, and have never changed over.

Simplified writing styles can be just as fluent and rhythmical as cursive writing with systematic practice continued until mature habits are established. Upper grade teachers who have taught handwriting as maturation of primary grade manuscript writing have found satisfaction in the ease with which children are taught and the excellent results achieved.

It is not necessary to maintain skill in two different styles, one for fluent writing, the other for the occasion when "printing" is better. With a slightly modified manuscript hand, it is entirely possible to write with good speed and still produce clean, neatly written script. There is no question that simplified styles have advantages, too, for adult illiterates and non-English speaking people learning to read and write English.

The school child's chief use of handwriting is for school papers and reports; the adult most commonly uses handwriting for his informal letters and memoranda. A clear-cut manuscript style is well suited for all these writing purposes.

Instead of making a complete change-over from manuscript to cursive letter forms, just a few slight adjustments are needed for maturation of the manuscript style writing learned in the primary grades. Mr. George Thompson, a British handwriting authority, has devised a slight-slant, slightly joined style which he refers to as "Italic script" This style is developed from the children's primary grade manuscript style writing without any basic changes in letter forms or strokes. Mr. Thompson has prepared a copy book with directions for developing fluent, rhythmical writing in this simplified style, writing that is fluent as well as legible and beautiful.[3] The pupil slants his paper when writing on a flat surface and may or may not slant the writing, depending on his personal inclinations.

Even though the quality of writing continues to be good, systematic and continuous practice is needed throughout the elementary grades to step up the rate of writing and to automatize rhythmical writing habits.

Simplified writing styles done with modern tools are by no means lacking in individuality. Each child's writing as he matures, increasingly mirrors his own personal adjustments and preferences with no essential difference in basic letter styles. At maturity, individual variability in the established writing hand is always apparent.

During the past thirty years or so, a period of rapid change in elementary education, experimental work in handwriting has been at a virtual standstill. Now with new-style handwriting tools and the frequent demand for clear-cut writing, a vigorous program of research is needed to appraise existing practices and to try out new methods. This experimentation should be carried out on a broad basis in all types of schools, and it should be accompanied by an educational program for parents to acquaint them with new practices in teaching handwriting.

FOOTNOTES

1. Teaching Handwriting, Curriculum Bulletin No. 3, 1960-61 (New York: Bureau of Curriculum Research, Division of Elementary Schools, Board of Education, 1961).
2. Gertrude Hildreth. "Manuscript Writing After Sixty Years," Elementary English, XXXVII (January 1960), pp. 3-13.
3. George Thompson. Better Writing, Puffin Picture Book, No. 96 (Middlesex, England: Penguin Books, Ltd., 1954).

Field News

COOPERATIVE RESEARCH PROGRAM

The Department of Health, Education and Welfare announces that some changes are being made in the review procedure of the Cooperative Research Program. All proposals will be divided into four groups: 1) those dealing with the learner as an individual (Personnel and Processes), 2) those dealing with the content of education (Curriculum and Methods), 3) those dealing with the environment as it influences the learning process (Environment), and 4) those dealing with areas being given special emphasis, such as "Project English," and involving both research and other activities.

The review procedure will be conducted in two stages. There will be a technical evaluation by Panels corresponding to the four areas. The Panels will select those proposals which met the criteria. The second stage will consist of a review of those recommendations by the Research Advisory Council which will be concerned with overall research policy and needs for research information.

It is expected that the Panels and the Council will meet only twice each year. September 1 and February 1 will be the deadline dates for submission of proposals. The Panels will meet in March and the Council in April. Investigators will be notified within two weeks after the Council meeting about the status of their proposals.

RELIABILITY OF WRITTEN COMPOSITIONS

To determine the standards competent readers use in judging freshmen compositions, Educational Testing Service has conducted a study under a grant by the Carnegie Corporation. The judges chosen for the study were 53 outstanding representatives of six different fields: college English teachers, social and natural scientists, writers and editors, lawyers and law professors, and business executives. Three hundred essays written by first year students at Cornell, Middlebury, and the University of Pennsylvania were used. The essays were written on the topics "Who Should Go to College?" and "When Should Teenagers be Treated as Adults?"

Each judge read the 300 essays on both topics and noted what he liked or disliked about each. The grades assigned by each reader were correlated with each other and with three College Board scores: SAT-V, SAT-M, and the English Composition Test. The resulting 56 × 56 correlation matrix was factored and rotated to oblique simple structure.

Six factors were identified. One factor was the usual "test" factor. Five other factors were identified as Ideas; Structure; Style, Interest, and Originality; Mechanics; and Effectiveness of Expression.

The factor analysis indicated that objective measures have moderately high correlations with the judgment factors identified as Mechanics (.50) and Effectiveness of Expression (.45), but very low correlations with the other three factors. The investigators conclude that tests can measure the former, but that "we know very little about how to test for Ideas; Structure; and Style, Interest and Originality — either by objective tests or by essays."

PHI DELTA KAPPA RESEARCH SYMPOSIA

Phi Delta Kappa announces that the next two research symposia will deal with "Research Problems Related to the Structure of Knowledge in Education" (University of Illinois, 1963) and "Selection, Recruitment, Motivation, and Preparation of Educational

Researchers” (1964, host institution has not yet been named).

REVISION OF ACE PSYCHOLOGICAL EXAMINATIONS AND ENGLISH COOPERATIVE TESTS

The Cooperative Test Division of Educational Testing Service states that it has very definite plans for the revision of the ACE Psychological Examination and tentative plans for the revision of the Cooperative English tests of usage, spelling and vocabulary.

RESEARCH AT THE UNIVERSITY OF KANSAS

Some dissertations completed or underway in the School of Education at the University of Kansas are as follows:

Robert R. Dunwell. “An Evaluation of Barnard's Statement of Administrative Theory Using Thompson's Criteria.”

James Jaret Lewis. “The Relationship of Aptitude and Prior Achievement to Performance at the Air Force Academy.”

Andrew Spencer Ely Martin. “Machine-Accounting Principles and Procedures as Applied to Pupil Census and Attendance Accounting in the Kansas City, Missouri, Public Schools.”

Lynn Robert Osborn. “An Analysis of the Relationship between Speech Performance and Performance on Written Examinations of Course Content in a Beginning College Speech Course as Reflected in Assigned Grades.”

Eldon Eugene Snyder. “The School as a Social System: A Study of Selected Aspects of a Student Social Structure.”

Raymond Arthur Urbanek. “A Study of Self-Scoring Tests in Terms of Learning Effects, Reliability and Psychological Effects.”

Maassouma M. Kazim. “Present Views Concerning the Nature of Mathematics and Their Implications for the Development of Criteria for Evaluating Secondary School Mathematics Curricula.”

Kenneth L. Curtis. “Follow-up Study of Persons Who Applied for and Were Granted Admission to the University of Kansas but Who Did Not Enroll.”

EMPHASIS ON ACADEMIC SUBJECTS AND PHYSICAL FITNESS

The Ohio Association of School Administrators has attempted to find out the effect of public clamor for more emphasis on academic subjects and on physical fitness. A study conducted by the Association indicated that a greater percentage of students were taking science, mathematics, foreign languages, social studies, English, and physical education. There was reduced emphasis in fine arts, practical arts, and driver education — the biggest percentage drop in enrollment was found to be 25.5 percent in fine arts for the period 1956-61.

* * * * *

Readers are invited to send items of research interest to John Schmid, Professor, College of Education, University of Arkansas, Fayetteville, Arkansas.

Among Those Present

REGINALD G. KIERSTEAD did his undergraduate work at Vermont and his graduate work at Columbia. He has held a number of school administration posts in Vermont and is now supervising principal at Laurens Central School, Laurens, New York.

ROBERT LLOYD DOCTER holds degrees from UCLA and USC. After eight years as a teacher and curriculum consultant in the Los Angeles schools, he is now assistant professor of education at San Fernando Valley State College.

GERTRUDE HOWELL HILDRETH is professor of education, Brooklyn College, The City University of New York. She is the author of Teaching Reading, Readiness for School Beginners, Learning the Three R's, and other publications in the language arts field.

MAURICE W. LAUFER is director of Bradley Hospital, member of the Institute for the Health Sciences, Brown University, and adjunct professor, Rhode Island College. He holds an MD from Long Island College of Medicine. ANTHONY DAVIDS is a Harvard PhD who is associate professor of psychology at Brown and chief psychologist at Bradley. MAX TALMADGE holds degrees from City College of New York and the University of Utah. He is senior psychologist at Bradley.

CARRIE M. SCOTT is remedial teacher in the Bend, Oregon, Public Schools. Previous experience includes assignments in Kansas and Colorado elementary schools, and on the staff of the University of Wichita. She holds BA and MA degrees from Colorado State.

Arvil S. Barr, Deceased, Editor and Publisher, 1928–1962

The Journal of Educational Research

Dedicated to the Scientific Study of Education

VOLUME 56 MARCH 1963 NUMBER 7

It Says in The JER 338

Factors Associated With the Progress of Students Enrolled in College Reading Program—J. Wesley Schneyer 340

Effects of Self-Directed Study on Course Achievement, Retention, and Curiosity—Donald E. Hovey, Howard E. Gruber, and Glenn Terrell 346

Student Perceptions of Group Work in Teacher Education Classes—Edmund Amidon and Arthur Blumberg 352

Lola Gets What Lola Wants from Supervision—Clifton M. Claye 358

Relationship of Selected Adjustment Factors, College Ability, and Achievement of Drop-Outs and Nondrop-Outs of College Freshmen—Warren J. Vorreyer 362

Teacher Roles Before and After Internship—Russell P. Kropp and John E. Anderson 366

The Effect of Visual Imagery Upon Spelling Performance—Leon D. Radaker 370

An Evaluation of the Vocational Counseling Program of the Employment Security Service in Kanawha County—William V. Wagner 373

A Longitudinal Study of Auditory Discrimination—Bertha Boya Thompson 376

Degree of Teacher Domination in Group Processes and Student Attitude Toward Instruction in Secondary Methods Classes—Kenneth H. Hoover 379

Evaluating Teaching Personality Before Student Teaching Begins—Lebern N. Miller 382

Teacher Enthusiasm—Victor E. Mastin 385

The Comprehension Subtest of the Wechsler Adult Intelligence Scale as an Indicator of Social Awareness—Robert M. Roth 387

Personality Needs of Academically Superior Students and Their Parents—Kurt Haas 389

Among Those Present 391

The Journal of Educational Research is published at Madison, Wisconsin, monthly 10 times a year (the May–June and July–August issues being combined). The subscription price is $7.50 a year, plus $1 postage for all subscriptions outside the continental United States. Single copies, $1. Second class postage paid at Madison, Wisconsin. Copyright 1963 by Dembar Educational Research Services, Inc. Address all correspondence care of DERS, Box 1605, Madison 1, Wisconsin.

Published by DEMBAR EDUCATIONAL RESEARCH SERVICES, Inc. WALTER FRAUTSCHI, President. Prof. WILSON B. THIEDE, Vice President and Publisher. Prof. CLARENCE A. SCHOENFELD, Assistant to the Publisher. ARNOLD CAUCUTT, Treasurer and Business Manager. JEAN WILLIAMS POTTER, Supervisor of Editorial Services.

Executive Editors

WILSON B. THIEDE, Professor of Education, Educational Administration, and Agricultural and Extension Education, The University of Wisconsin, Madison, Wisconsin. Chairman.
THEODORE HARRIS, Professor of Educational Psychology and Education, The University of Wisconsin, Madison, Wisconsin.
L. JOSEPH LINS, Professor and Coordinator of Institutional Studies, The University of Wisconsin, Madison, Wisconsin.
PHILIP LAMBERT, Professor of Educational Psychology and Educational Administration, The University of Wisconsin, Madison, Wisconsin.
JOHN SCHMID, Professor of Education and Chairman of the Department of Educational Foundations, University of Arkansas, Fayetteville, Arkansas.

Editorial Board

JACOB O. BACH, Head, Educational Research Bureau, Southern Illinois University, Carbondale, Illinois.
WILLIAM A. BROWNELL, Dean, School of Education, University of California, Berkeley, California.
ROBERT A. DAVIS, Professor of Educational Research, George Peabody College for Teachers, Nashville, Tennessee.
GUY D. McGRATH, Dean, College of Education, Arizona State University, Tempe, Arizona.
HAROLD E. MITZEL, Director, Office of Research and Evaluation, Division of Teacher Education, College of the City of New York, New York City.
WILLARD C. OLSON, Dean, College of Education, University of Michigan, Ann Arbor, Michigan.
ARTHUR E. TRAXLER, Executive Director, Educational Records Bureau, New York 32, New York.

Associate Editors

HAROLD M. ANDERSON, Associate Professor of Education, University of Colorado, Boulder, Colorado.
VERNON E. ANDERSON, Dean, College of Education, University of Maryland, College Park, Maryland.
RUDYARD, K. BENT, Professor of Education, University of Arkansas, Fayetteville, Arkansas.
EMMETT ALBERT BETTS, Research Professor, University of Miami, Coral Gables, Florida.
LAWRENCE P. BLUM, Assistant Director, Student Personnel Services, University of Wisconsin-Milwaukee, Milwaukee 3, Wisconsin.
J. MURRAY LEE, Chairman, Department of Elementary Education, Southern Illinois University, Carbondale, Ill.
D. WELTY LEFEVER, Professor of Education, University of Southern Caifornia, Los Angeles, California.
S. E. TORSTEN LUND, Associate Professor of Education, University of California, Berkeley, California.
WILLIAM A. McCALL, Lecturer in Education, University of Miami, Coral Gables, Florida.
DELYTE W. MORRIS, President, Southern Illinois University, Carbondale, Illinois.
VALWORTH PLUMB, Chairman, Division of Education and Psychology, University of Minnesota (Duluth Branch), Duluth, Minnesota.
DAVID G. RYANS, Rand Corporation, Santa Monica, California.
LESTER B. SANDS, Professor of Education, University of California, Santa Barbara, Goleta, California.
MANDEL SHERMAN, President, and Consultant, Emotional Health Association, Los Angeles, California.
MILDRED B. STANTON, Consultant, Special Education, State Department of Education, Hartford 15, Conn.
W. W. THEISEN, Professor of Education, Marquette University, Miwaukee, Wisconsin.
M. R. TRABUE, Professor of Higher Education, University of Kentucky, Lexington, Kentucky.
J. WAYNE WRIGHTSTONE, Director, Bureau of Educational Research, Board of Education of City of New York, Brooklyn, New York.

Editors Emeriti

E. J. ASHBAUGH, Dean Emeritus, School of Education, Miami University, Oxford MR50, Ohio.
HARRY J. BAKER, 19050 Wiltshire, Lathrup Village, Michigan.
FOWLER D. BROOKS, Director Emeritus, Department of Psychology, DePauw University, Greencastle, Ind.
LEO J. BRUECKNER, Professor Emeritus, University of Minnesota, Minneapolis, Minnesota.
B. R. BUCKINGHAM, 33 Chatham Circle, Wellesley Hills, Massachusetts.
EDGAR A. DOLL, Director Emeritus, Devereux Schools, Devon, Pennsylvania.
HARL R. DOUGLASS, Director Emeritus, School of Education, University of Colorado, Boulder, Colorado.
E. DUNCAN GRIZZELL, Dean Emeritus, University of Pennsylvania, 640 Maxwelton Court, Lexington, Ky.
NEST HORN, Professor Emeritus of Education, State University of Iowa, W307 East Hall, Iowa City, Iowa.
RL HUDELSON, Professor Emeritus of Education, West Virginia University, Morgantown, West Virginia.
A. M. JORDAN, Professor Emeritus, University of North Carolina, Chapel Hill, North Carolina.
ALFRED S. LEWERENZ, Director Emeritus of Evaluation and Research, 4107 Holly Knoll Drive, Los Angeles, California.
A. R. MEAD, Director Emeritus of Educational Research, University of Florida, 1719 N. W. 6th Ave., Gainesville, Florida.
M. J. NELSON, Dean Emeritus, Iowa State Teachers College, 6033 West Orange Drive, Glendale, Arizona.
JESSE B. SEARS, Professor Emeritus, Stanford University, Stanford, California.
CARLETON W. WASHBURNE, Chairman Emeritus, Department of Education, [illegible] Brooklyn, New York.

It Says in The JER...

WHAT IS BEING DONE NOW, in the teacher preparation institutions of the nation, to identify and redirect students whose personality traits seem to limit their chances of success in teaching? After surveying 30 colleges and universities, Miller finds a great diversity of screening procedures and infrequent use of standardized tests.

CAN YOU PREDICT college success? Vorreyer offers formulae for identifying potential drop-outs. No single item of information available on freshmen affords an adequate index, he says. You need a constellation of selected traits and qualities to make a valid prediction.

WHAT FACTORS ARE ASSOCIATED with student improvement in a college remedial reading course? According to a Schneyer study, high-improvers tend to be students who have made tentative occupational choices, who read serious books, and who rate their trouble as an inability to concentrate. Poor-progress students tend to show a combination of low motivation and high anxiety. But Schneyer says his major conclusion was that he could find no instrument capable of discriminating sufficiently among students to permit prediction of progress on an individual basis.

CHILDREN RECEIVING TRAINING in the creation of images will score higher in spelling, Leon Radaker reports from a Pennsylvania study. He thinks imagery practice should improve reading skills, too.

AUDITORY DISCRIMINATION and intelligence are highly correlated with success in primary reading, according to data presented by Thompson from an Ohio study.

STUDENT INTERNS discover the actual role of a teacher to be much like their conception of the ideal role of a teacher, Kropp and Anderson report. Their directing teachers and methods teachers are in only moderate agreement as to the ideal role of the teacher, however, and in even less agreement about the actual role of the teacher.

WHAT EFFECT do different degrees of teacher-centeredness have upon student attitudes toward the instructor and the course? Hoover could detect none at all in an Arizona study.

IN A DEPTH ANALYSIS of employment counseling services in Kanawha County, West Virginia schools, Wagner discovers major strengths and weaknesses and makes recommendations for improvement.

ARE BOYS AND GIRLS likely to learn more under enthusiastic teachers than under teachers who appear indifferent? Mastin discovers that teacher enthusiasm has a real influence on pupil learning and pupil attitudes.

UNDER THE INTRIGUING TITLE, "Lola Gets What Lola Wants from Supervision," Claye reminds us that unless a teacher desires change, no permanent change will be effected. He believes that teachers want to improve the teaching-learning situation and expect principals and supervisors to lend a hand.

SUPERIOR STUDENTS AND THEIR PARENTS do not exhibit distinctly unusual personality needs, Haas reports. They are successful "because they have more ability and apply it more diligently."

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 7, March 1963)

Factors Associated With the Progress of Students Enrolled in a College Reading Program*

J. WESLEY SCHNEYER
University of Pennsylvania

The Problem

IT WAS THE PURPOSE of the study here reported to investigate the factors which might distinguish students who made progress in a college reading improvement course from students who failed to make progress. This investigation was undertaken in an attempt to assist those who are responsible for selecting students with the greatest potential for benefiting from a reading and study skills course. By awareness of the characteristics associated with students who made large improvement in a course, the instructor of the reading program would have available a set of criteria for selecting those most able to achieve gains in the program. On the other hand, if those students who would be least likely to gain from a course could be identified, and the characteristics associated with lack of improvement of their reading and study skills spelled out, then perhaps more appropriate instructional procedures, or supplementary therapy could be provided to enable even the potentially low improvers to increase the amount of gain they might expect.

Procedures

Of the 91 students who enrolled for group instruction in the University of Pennsylvania's Reading Clinic Program during the academic year 1958-59, a total of 71 students remained to complete the Program, including the pre-Program and post-Program testing. These 71 students compose the sample for the present investigation. However, certain initial data were available for the students who enrolled for the Program but who did not return for the final testing. Data concerning this group, called the Attrition Group, are presented where pertinent.

Each of the students who enrolled for the Program was administered the *Cooperative Reading Test* (4) before and after the Program. The difference between the pre- and post-total reading test scores was the criterion used for determining the amount of progress achieved in the Program.

Based on the changes between the initial and final *Cooperative Reading Test* total scores, three reading progress groups were formed: A high and Low Improvement group, with 21 students in each group, and an Average Improvement group, with 29 students.

A battery of reading and study skills tests, personality inventories, and a personal interview were administered to each student in the Program. Two of the reading tests were administered before and after the Program, while the study habits and personality inventories were administered only at the beginning.

The Reading and Study Skills Program consisted of a seven weeks course for all students, who met in groups of ten for two classroom sessions and one laboratory session per week. During the classroom sessions, certain reading and study skills were taught and practiced by the students, while practice in increasing speed of reading was provided during the laboratory period.

In addition to the battery of tests and the interview, data concerning scholastic aptitude and academic achievement were obtained for each student from his school office.

The mean differences of the various test scores for the three reading progress groups were tested for statistical significance by analysis of variance techniques. Differences on the questionnaire items were tested for significance by chi square. Differences were considered statistically significant if they were at the .05 level or higher.

Findings

Scholastic aptitude. Scores for each student on the College Entrance Examination Board's *Scholastic Aptitude Test* (2) were available from the school office. The *Scholastic Aptitude Test* provides three scores: a verbal aptitude score (VAT); a mathematical aptitude score (MAT); and a total scholastic aptitude score (SAT).

The analysis of variance results of the VAT scores for the three progress groups yielded an F ratio that was significant (.05>P>.02). Further analysis of the differences by Snedecor's modification of the Tukey test indicated that all of the mean differences were significant. The High Improvement group had a significantly lower verbal ability mean score than either the Average or the Low Improvement groups. (See Table I.) The Low Improvement group obtained a significantly higher VAT mean score than either of the other groups. The mean VAT score for the Attrition group was only 13.93 higher than that of the High Improvement group and 49.59 below that of the Low Improvement group. A negative correlation of -.26 was found between VAT scores and progress in the reading Program as measured by the criterion factor. This correlation was significant at the .05 level.

---

*An abstract of the author's Ed. D. dissertation completed at the School of Education, University of Pennsylvania, January 1961.

Analysis of variance of the mean differences of the MAT scores for the three progress groups revealed no statistically significant differences in mathematical ability. There was a tendency for the High Improvers to have a higher MAT score than the Low Improvers. The correlation between MAT scores and gain in the Program resulted in an r of .04, which was not significant. The Attrition group achieved a mean MAT score of 565.40, which was the highest mean score for all of the groups.

The analysis of variance of the SAT scores for the three reading progress groups yielded no statistically significant differences. The correlation between SAT scores and progress in the Program was .01, which was not significant. There was a tendency for the Low Improvers to achieve the higher SAT score. The SAT score for the Attrition group was similar to that of the Average Improvers, which was in between that of the High and Low Improvement groups.

The analysis of variance revealed no significant differences on MAT minus VAT differences for the three progress groups. The mean difference of the High Improvers was 82.81 compared to 5.4 for the Low Improvers. This difference did approach significance at the .05 level, however. The mean difference for the Attrition group was 82.95, second only to the highest of all the groups. The correlation between MAT minus VAT differences and gain in the Program was .28, which was significant at the .02 level of confidence.

Reading and study skills. This part of the investigation was concerned with the relationships of certain reading and study skills to progress in the Program. The questions of concern were: Is the ability to read and understand different kinds of material, fiction and factual, related to progress in the Program? Would the students who made the greatest gains in the reading rate and understanding of story-type and of factual material be the same students who made the greatest progress in the Program?

Three tests were administered to each student enrolled in the Program, both before and after the Program, to obtain data relevant to this area. The Diagnostic Reading Test (Survey Section), Form A, (3) as a pre-test and Form B as a post-test were administered to obtain data related to rate and comprehension of story-type material. Two articles from the SRA Better Reading Book 3, (6) one as a pre-test and the other as a post-test, were administered to obtain data relevant to the student's rate and comprehension of factual type material. The "Informal Reading and Study Test," developed especially for this investigation by the investigator, was administered in its single form both before and after the Program as a measure of the student's ability to read, take notes, comprehend, study, and master a chapter of textbook type material.

The analysis of variance of the reading rate scores on the Diagnostic Reading Test for either pre-test, post-test, or gains in rate for the three reading progress groups yielded no significant differences. The High Improvers tended to have the highest initial reading rate, the highest final reading rate, but to achieve the smallest gain in rate. The Low Improvement group realized the greatest gain in reading rate of story-type material.

The mean initial reading rate for the Attrition group was the lowest initial rate among the reading groups.

The analysis of variance of the initial, final and gain comprehension scores of the Diagnostic Reading Test (Survey Section) yielded no significant differences for the three reading improvement groups. The comprehension scores of the High Improvement group were lower than the Low Improvement group at the beginning of the Program, but were higher at the end of the Program and indicated the greater gain in comprehension of story-type material. All of the F ratios, however, were less than 1.00. The mean comprehension score for the Attrition group was the lowest initial reading comprehension score for any of the reading groups.

The analysis of variance of the initial, final and gain scores in reading rate of the articles in the SRA Better Reading Book 3 (6) revealed no significant differences for the three reading progress groups. All of the F ratios were less than 1.00. There was a tendency for High Improvers to begin the Program with the lowest initial reading rate, to end the Program with the lowest rate, but to make the greatest gain in reading rate of factual material. No data were available for the Attrition group.

The analysis of variance of the initial, final and gain scores in comprehension of factual material yielded no significant differences among the three reading improvement groups. There was a tendency for the High Improvers to begin with a higher comprehension score, to end with a lower comprehension score, but to make greater gains in comprehension than the Low Improvement group. The Lower Improvement group achieved a slightly negative difference between their initial and final comprehension scores. This difference approached

TABLE I

RANGES, MEANS, AND STANDARD DEVIATIONS OF SCHOLASTIC APTITUDE TEST SCORES FOR HIGH, AVERAGE, LOW IMPROVEMENT AND ATTRITION GROUPS

Test	Statistic	High Improvement Group	Average Improvement Group	Low Improvement Group	Attrition Group
VAT	Range	377-611	329-743	375-638	405-512
	Mean	469.52	498.21	532.04	482-45
	S. D.	80.27	68.02	66.33	42.65
MAT	Range	427-765	328-800	345-724	416-669
	Mean	552.33	539.27	537.47	565.40
	S. D.	77.49	96.94	96.29	68.50
SAT	Range	868-1370	760-1380	857-1362	845-1183
	Mean	1024.09	1031.96	1063.19	1035.80
	S. D.	120.03	239.83	145.01	222.86
MAT minus VAT	Range	-57 to +247	-192 to +220	-253 to +199	+10 to +211
	Mean	82.81	41.07	5.38	82.95
	S. D.	79.70	97.27	105.10	55.42

significance at the .05 level. Data were not available for the Attrition group.

The analysis of variance of the initial, final and gain scores for underlining or note-taking on the "Informal Reading and Study Test" revealed no significant differences among the three reading improvement groups. The High Improvers tended to achieve a higher initial score, and a lower final score than the Low Improvers. The Low Improvers made greater gains in underlining ability than did the Higher Improvers. None of the F ratios, however, were greater than 1.00.

The analysis of variance of the initial, final and gain scores in summarizing on the "Informal Reading and Study Test" (which was developed by the investigator especially for this study) yielded no significant differences among the three reading improvement groups. There was a tendency for the High Improvers to have lower initial and final scores than the Low Improvers, but for the High Improvers to achieve a greater gain in summarizing ability. None of the differences was significant.

Study habits and attitudes. Study habits and motivation for college work have been shown to be related to both academic achievement and growth in reading skills. The questions of concern here were: Are study habits, methods of study, and attitudes or motivation for study, related to progress in the Program? Would the students making the greatest progress in the Program tend to have the greatest motivation for study and the best methods and study habits?

To obtain data related to these questions, the Study Habits Checklist (5) and the Brown-Holtzman Survey of Study Habits and Attitudes (1) were administered to each student. A total score was obtained for the Study Habits Checklist, while separate "study-mechanics" and "study-attitudes" scores as well as total study habits scores were obtained from the SSHA.

The analysis of variance of the Study Habits Checklist scores revealed that there were no significant differences in the study habits of the three reading progress groups. The differences did, however, approach significance at the .10 level of confidence. The tendency was for the Low Improvers to have better study habits. The mean study habits score for the Attrition group was the lowest of any of the groups. The correlation between Study Habits Checklist scores and growth in the Program was -.29 which was significant at the .02 level.

The analysis of variance of the "study-attitude" scores, the "study-mechanics" scores, and the total scores of the SSHA yielded no statistically significant differences. The High Improvers tended to have better "study-mechanics" and total study habits than the Low Improvers. The F ratios for these scores were less than 1.00. However, the High Improvers also had better "study-attitudes" or motivation toward study than did the Low Improvers. This difference resulted in an F ratio of 2.76 which approached significance at the .05 level of confidence. The correlation between "study-mechanics", "study-attitudes", and total study habits scores and gains in the Program, were all negligible and not significant. The mean total study habits score, "study-mechanics" score, and "study-attitudes" score, for the Attrition group were all lower than those of the other reading groups.

Ranking troublesome reading and study problems. It was noted early in the investigation that most students who enrolled in the Program or who came for conferences with the investigator, tended to believe that their chief source of reading study difficulty was a lack of reading speed. To see if this tendency to consider speed important was related to growth in speed of comprehension, students enrolling in the Program were asked to rank a list of seven common reading-study problems in the order in which the student considered them to be important.

The hypothesis was tested by correlating the students' ranking of the importance of gains in speed of reading to the student with his actual gain as measured by changes in speed of comprehension scores on the Cooperative Reading Test. A separate correlation was computed for the High Improvers, Average Improvers, and for the Low Improvers. The correlation between ranking and gain in speed of comprehension was for the High Improvers: .48 (significant at better than .05 level); for the Average Improvers: .36 (not significant); and for the Low Improvers: .32 (not significant). The apparent motivation of the High Improvers to make gains in reading rate seemed to play a greater part in their growth in speed of comprehension, than it did for either of the other groups.

Trouble with concentration was the only problem that was ranked important to a significantly greater extent than it was ranked unimportant. A significantly greater number of High Improvers ranked this problem important than ranked it unimportant (P=.005). There was no significant difference in the ranking of this problem by the Low Improvers.

Academic achievement. The final grade-point average for each student enrolled in the Program was obtained from his school office at the end of the academic year following his participation in the Program.

The analysis of variance of the grade-point averages for the three reading progress groups, indicated that the differences among the High, Average, and Low Improvement groups were not statistically significant. The F ratio was less than 1.00. The mean of the High Improvers was 2.98 which was .23 above that of the Low Improvers and .33 above that of the Attrition group, which had the lowest grade-point average of all the groups.

Personal, educational, and social factors. Each student who came to the Reading Clinic for a conference was asked to complete a questionnaire which contained a number of items relative to his personal, social, and educational background. These data were available for each student enrolled in the Program. Thirty-two items were selected from the questionnaire for analysis and comparison between the High Improvers and Low Improvers. The difference between the relative frequencies of responses for each item for the High and Low Improvers was tested for significance by chi square.

A difference which was significant approaching the .01 level of confidence was found between the types of books read most frequently by the High Improvers and those read by Low Improvers. The High Improvers tended to read theoretical, academic, and classical books more frequently than they read popular books and current event books. This preference was reversed for the Low Improvers.

A difference approaching significance at the .05 level of confidence was found in the tentative selection of a vocational choice at the time of enrollment in college between the High Improvers and Low Improvers. A larger number of High Improvers than Low Improvers had made a tentative choice of occupation at the time they had enrolled in college.

There was a tendency, but not significant, for more of the fathers of High Improvers to have had at least one year of college than for the fathers of Low Improvers. No difference was found between the education of mothers of the two groups.

None of the other factors was significant, approached significance, nor revealed definite trends.

Personality factors. Research had suggested that one of the factors related to growth in reading ability might be personality. Two instruments were used in the assessment of personality. The SA-S Senior Scale (7) evaluated students on the basis of emotional stability and permeability. The Inventory of Beliefs (8) evaluated students on the basis of their tendency to accept or to reject stereotyped attitudes.

The analysis of variance of the SA scores of the SA-S Senior Scale revealed no significant differences among the High, Average, or Low Improvement groups. The F ratio was less than 1.00. There was a small tendency for the High Improvers to be more permeable than the Low Improvers. The Attrition group had a mean SA score indicating that this group was the most permeable of all the groups.

Analysis of variance of the S scores of the SA-S Senior Scale for the High, Average, and Low Improvement groups revealed no significant differences. The F ratio was 1.55. A comparison of the mean scores of the three groups indicated that the High Improvers were more stable than the Low Improvers. The lower mean score of the Low Improvement group indicated a tendency for this group to be more anxious than the High Improvement group. However, the differences were not significant. The Attrition group had a mean S score which indicated that this group was the most anxious of all the groups.

Chi square tests of the relative frequencies of the four personality types, and various combinations, revealed no statistically significant differences. High Improvers tended to be students who were both permeable and anxious. When the groups were combined for each of the separate factors, 66.6 percent of the High Improvers were permeable, while only 33.4 percent were impermeable. Of the 21 Low Improvers, only 38.1 percent were permeable, while 61.9 percent were impermeable. This difference approached significance at the .05 level.

When the personality types were combined on the basis of similarity in emotional stability, 47.6 percent of the High Improvers were emotionally stable, while 52.4 percent were anxious; for the Low Improvers, however, only 23.8 percent were stable, while 76.2 percent were anxious. This difference approached the .10 level of significance. The Attrition group was not included in these combinations.

The analysis of variance of the Inventory of Beliefs scores for the High, Average, and Low Improvers, revealed no statistically significant differences. The F ratio was 1.8. Inspection of the mean scores for the three reading groups revealed that there was a greater tendency, though not significant, for the Low Improvers to accept stereotyped statements. The mean for the Attrition group indicated that these students had a greater tendency than any of the groups to accept stereotyped ideas.

Conclusions

A major conclusion to be drawn from this investigation is that none of the instruments employed in the study (reading tests, personality inventories, or study habits inventories) was able to discriminate among students taking a reading improvement course with sufficient statistical significance to permit prediction of progress on an individual basis. There is an urgent need to develop tests which contain items that are not transparent and to which responses cannot be falsified to enable the testee to respond in a manner which will produce the highest score. This is true for personality inventories and instruments for evaluating study habits and attitudes toward learning.

However, it should be noted, that while the present instruments were unable to discriminate with sufficient reliability for individual prediction, certain characteristics and trends were observed which would seem to indicate tendencies that would describe High Improvers and Low Improvers as a group.

Insofar as the results of the present study are concerned, prediction of success in a reading course can be based on certain tendencies, some of which are significant, others which approach significance, and some of which are not significant, but which, however, appear to indicate trends which may be observed in the data and are supported by the investigator's own experience.

The significant tendencies are for the High Improvers to be students with low verbal aptitude scores; to be students who make tentative occupational choices at the time of entrance to college; to be students who read books of an academic, classical, and serious nature; and to be students who rank trouble in concentration as a major reading-study difficulty.

Those tendencies which approach significance are for the High Improvers to be students whose mathematical aptitude scores exceed their verbal aptitude scores; to be students who are highly flexible persons; and to be students who are well motivated for academic studies.

Those tendencies which are not significant but which the investigator believes further study may reveal to be significant are for High Improvers to be students whose initial reading scores are low; and to be students who are emotionally stable and non-anxious.

An important factor noted for the Attrition group was that their motivation toward learning as measured by the "study-attitudes" scores from the SSHA were the lowest of any of the reading groups. The students who failed to return for the final tests would appear then, to be the least motivated of the students in any of the groups. In addition, as revealed by the SA-S data, these students were also the most anxious. The combination of low motivation and high anxiety appeared to be an important factor in influencing students who failed to return for final tests. Other factors, unknown at present, may have influenced these students. For example, they may not have been satisfied with their progress in the Program, and the pressure of studying for final examinations combined with their extreme anxiety might have been enough to prevent them from returning for the re-tests. While final answers to the question of why the students in the Attrition group did not return to complete their final tests will never be known, the data concerning their performance on the SA-S test and on the SSHA (high anxiety combined with low motivation) suggests that students with extreme scores on these tests would benefit from the individualized attention.

One of the implications of this study which merits attention is the observed tendency of High Improvers to achieve low scores on the verbal subtest of the Scholastic Aptitude Test. Correlations ranging from .58 to .70 between VAT scores and Cooperative Reading Test scores indicate that the two tests are measuring similar skills to some extent. It seems probable that some of the students who obtained low VAT scores did so because of limitations in their reading ability. It may well be that students whose low verbal aptitude scores are apparently due to poor reading ability may improve their verbal aptitude scores after a period of intensive instruction in reading. Further research in this area would seem to be warranted.

REFERENCES

1. Brown, William F., and Wayne H. Holtzman. Survey of Study Habits and Attitudes, New York: The Psychological Corporation, 1953.
2. College Entrance Examination Board, Scholastic Aptitude Test, Princeton: Educational Testing Service, 1951.
3. Committee on Diagnostic Reading Tests, Diagnostic Reading Tests, Survey Section, Chicago: Science Research Associates, Inc., 1947.
4. Davis, Frederick and others. Cooperative English Tests, C2: Reading Comprehension, Princeton: Cooperative Test Division, Educational Testing Service, 1951.
5. Preston, Ralph C., and Morton Botel, Study Habits Checklist, Chicago: Science Research Associates, Inc., 1957.
6. Simpson, Elizabeth, SRA Better Reading Book 3, Chicago: Science Research Associates, Inc., 1951.
7. Smith, Donald E. P., and Wood, Roger L. SA-S Senior Scale, Bureau of Psychological Services, University of Michigan, 1956.
8. Stern, George. Inventory of Beliefs, Syracuse University, (mimeographed).

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 7, March 1963)

Effects of Self-Directed Study on Course Achievement, Retention, and Curiosity*

DONALD E. HOVEY
United States Air Force Academy
HOWARD E. GRUBER and GLENN TERRELL
University of Colorado

RESEARCH ON college teaching methods has often been concerned with the student's degree of active involvement in the learning process. In such studies, the independent variable is the degree of activity required of the student, as in experiments comparing the lecture method with the discussion method. The dependent variable is usually some measure of the student's mastery of the material of the course. The frequent failure to obtain clear-cut differences in the effects of different treatments may be due to the choice of criterion measures lacking an explicit and well-founded relation to the treatment variable.

There is, for example, no obvious reason to expect independent study or other relatively unstructured learning situations to produce greater mastery of a standardized body of knowledge such as is usually tested in final examinations. On the other hand, it does seem reasonable to expect that if greater interest and curiosity about a given field is generated by independent study, it may mediate continuing efforts to solve problems and seek information in that field. In their turn, such continuing efforts may result in superior mastery of subject matter knowledge at intervals remote from the initial learning period. Indeed, this was the result obtained by Kersh (4) in a laboratory study of the "independent discovery" method.

Thus, curiosity as a directive aspect of cognitive activity assumes special significance in research concerned with long-range effects of educational methods. Unfortunately, curiosity has been given little direct attention in educational research. In some of the older studies cited by Stroud (6), and in Berlyne's more recent work (1), the occurrence and timing of questions proposed by the teacher or the textbook has been used as the independent variable, with mastery of subject matter as the dependent variable, but question-raising behavior has never been treated as a dependent variable.

In the present study, the independent variable is the degree of student activity elicited by the method of instruction: self-directed study (SDS) is compared with a conventional lecture method. In SDS, the number of lectures was reduced to one-third the number given in the control course. Thus, SDS differs from the method commonly called "independent study" in that SDS preserves the conventional course structure, in which one teacher guides many students through the same body of knowledge, whereas independent study typically involves a one-teacher:one-student relationship. As practiced in the present experiment, SDS differed from experiments in which discussion sections are substituted for lectures in that the SDS group attended only one-third as many formal class meetings with the instructor present as did the control group.

The dependent variables included the usual measures of mastery of course materials. In addition, measures of this kind were repeated after a delay of one year. But the most important departure from previous research was the attempt to measure curiosity as a dependent variable.

---

*We thank James Selkin for his help in collecting the data of the first phase of this experiment, and Dr. Morris Weitman for many valuable suggestions. The first phase of this study was supported by a grant from the Graduate Council on Research and Creative Work at the University of Colorado. The follow-up study is part of an extensive investigation of self-directed study at the University of Colorado, directed by H. E. Gruber. Financial support by the Fund for the Advancement of Education is gratefully acknowledged. In part, the research reported herein was performed pursuant to a contract with the United States Office of Education, Department of Health, Education, and Welfare.

Hovey (3) has conceived of curiosity as arising from the existence of gaps or inconsistencies in cognitive systems. These gaps occur when two cognitive organizations are compared (at various levels of awareness) and are found to be neither wholly congruent nor wholly incongruent.

Several hypotheses concerning the conditions under which gaps, and thereby curiosity behaviors, would occur were derived from this conceptualization. In brief, the development of cognitive gaps increases as some function of the complexity, number, variety, and relevance-to-self of cognitive systems employed by the individual in dealing with new experiences. The present study is concerned only with the last of these hypotheses. It is hypothesized that gaps are more apt to occur in materials associated with relevant than with irrelevant cognitive systems. By "relevant" is meant more central to the individual's concept of himself, of reality, and of his future. Cognitive systems become relevant to the degree that the individual is required to use them actively in adapting to the situations he confronts. Finally, the intervening variable, cognitive gaps, is coordinated to a behavioral variable: the questions raised by a student. Thus, the number and quality of questions generated by a student, either spontaneously or in response to the experimenter's request to ask questions, was our main behavioral index of curiosity.

An opportunity to test this formulation came to hand as an extension of a self-directed study experiment involving a course in educational psychology. The appropriateness of this situation for testing our hypothesis rests on the assumption that the S in the SDS condition was required more than the S in the lecture condition actively to use and manipulate the course materials, both in his individual study and in interacting with other members of his discussion group. Bloom's (2) research on students' thought-processes in the classroom lends support to our assumption. He found that during lectures, students' thought-processes characteristically take the form of passive recognition and cognizing of the material presented, whereas in discussion groups, there is more thinking concerned with problem solution and more comparison of one's own understanding of the material with statements made by others. In addition to measures of question-raising behavior one year after the course, we were interested in the S's direct expressions of interest and in overt behavioral tendencies to seek more information about the field covered in the experimental course.

Method

The Ss were two sections of students enrolled in Educational Psychology at the University of Colorado during the spring semester of 1957. Both sections were taught by one of the writers. By the flip of a coin it was determined that Section 1 (62 Ss scheduled to meet at 9:00 a. m. on Monday, Wednesday and Friday), would be taught by a lecture method. Section 2 (53 Ss originally scheduled to meet at 11:00 a. m. on Monday, Wednesday and Friday), would be taught instead by the method of self-directed study in small groups. At the University of Colorado, 9 o'clock and 11 o'clock sections are generally considered to be equally desirable, and there is no reason to expect students in these sections to differ systematically.

As entering freshmen, all Ss had received the University of Colorado Academic Achievement Test (AAT), a placement test giving separate and composite scores in English and mathematics. All Ss received the Otis Intelligence Test at the end of the semester in which the experiment was conducted. Near the beginning of that semester, all Ss received a test of Need for Academic Achievement (NacAch), based on the Liverant test (5). The lecture group, Section 1, was slightly superior on the AAT and on the Otis, while the SDS group, Section 2, was slightly superior on the NacAch scale. None of these differences was statistically significant.

To test our hypothesis about curiosity, as well as to assess the retention of course material, as many as possible of the original Ss were administered a questionnaire and a subject-matter test during the spring semester of 1958, about 10 months after the end of the course. Of the original 62 Ss in the lecture group, 40 were still enrolled at the University, and 36 of these took the post-test. Of the original 53 Ss in the SDS group, 35 were still enrolled and 30 took the post-test. The group obtained for the post-test were similar to the groups from which they stemmed with respect to scores on the Otis, the AAT, the NacAch scale, and final course grades.

Procedure

Although questions from Ss were permited in the lecture class, at least three-fourths of each period was devoted to a planned lecture, which the instructor consciously made more formal than is usual for him.

The SDS class was informed at its first meeting that it had been selected to try out a new method of teaching. Instruction in small self-directed groups was explained and the Ss then received one week of training in the application of techniques which had been found successful in promoting good communication in small groups. Work sheets were distributed which contained concrete suggestions as to the kind of behaviors which were successful in groups and discussion of the project was encouraged in class. In addition, demonstrations of successful and unsuccessful group processes were conducted for this class by several graduate students.

After this first week of training in group discussion technique, the SDS class was divided into ten small groups, each having five or six members. Each group was assigned a small room in which to meet each Monday and Wednesday at 11:00 a.m. The groups were equated with respect to AAT and NacAch scores.

The Friday meeting of the SDS class, at which time the instructor met with the entire class, was devoted to a discussion of common problems in group functioning and to questions concerning course content, particularly the latter. Although there was some initial awkwardness, the small discussion groups were meeting smoothly by the end of the second week of the term.

Topics for group discussion were supplied by the course instructor in the form of instruction sheets. These contained a list of the assigned readings, plus supplementary readings for the course. In addition, the instruction sheets contained an outline of the topics to be discussed, and a number of relevant discussion questions. It was suggested that students center their discussions around the instruction sheets. Throughout the semester, new sheets for the next unit of study were distributed to students in the experimental group approximately every two weeks.

When the groups first began meeting, leaders were assigned to each group. These leaders were student members of the group, selected by the course instructor. It was suggested that they keep group attendance records and take the role of leader in group discussions. It was further suggested that the leader's role be rotated throughout the group. In general, groups tended to drop the leadership suggestion after an initial tryout. One or two groups did retain a permanent leader, and most of the groups had the same person fill out the attendance slips from week to week. Students in most of the groups, however, tended to prepare and conduct group discussion on a basis other than appointed leadership.

Measures of Mastery of Course Material

The final examination consisted of an essay part and an objective part. The course grade was based on the final examination and on tests given during the semester.

The post-test, administered ten months after the course had ended, assessed retention of course content by means of 15 multiple-choice questions, 24 true-false questions, ten identification (short answer) questions, and a ten-item matching question. The multiple-choice and true-false questions were taken verbatim from the course final and other course examinations. The identification and matching questions were modified versions of questions which had appeared on course examinations.

Measures of Curiosity

A post-test questionnaire was administered which included three groups of items related to three aspects of curiosity. The first, and most important, was a group of five items dealing with question-raising behavior. In these five items, the S was asked, 1) to list as many questions as he could think of in the field of educational psychology, 2) comment on a paragraph about an educational problem, 3) list the problems or questions suggested by the previous paragraph, 4) list the things in the field of educational psychology that he would like to find out more about, and 5) list all the questions suggested by another paragraph dealing with an educational problem. Each item yielded a score for the number of questions raised, and the set of five items yielded a score for total number of questions raised. In addition, the protocol was scored for the merit and diversity of the questions raised. During the scoring, the scorer was kept ignorant of the experimental groups to which the protocols belonged.

There were five items in the second group, dealing with interest in the field of educational psychology. Two of these were Likert-type rating scales and three were open-ended questions which were later content-analyzed for expressions of interest in topics related to educational psychology. Again, the scorer was kept ignorant of the groups to which the protocols belonged.

There were five items in the third group, dealing with tendencies toward overt action to get more information about the field of educational psychology. Two of these were self-reports of reading actually done or other actions actually taken, and two dealt with hypothetical efforts the S would be willing to make. The final "action-indicator" was whether or not the S troubled to self-address a postcard asking for information about the outcome of the investigation. These "action-indicators" were used as measures of interest in educational psychology.

Results

At the end of the semester during which the course was taught, there was no significant difference between the groups on measures of mastery of course material. The SDS group was slightly but insignificantly superior to the lecture group on the final examination and in the final course grade. A similar result was obtained in the post-test ten months later. On the post-test, the SDS group was slightly superior in all three parts of the test of retention of course material. None of these differences was significant.

Question-raising Behavior—As shown in Table I, in each of the five items dealing with question-raising behavior, the SDS group was superior in number of questions raised. None of the differences on individual items was statistically significant. Since the intercorrelations among the scores on these items were negligible, it is reasonable to treat these five scores as independent observations.

TABLE I

NUMBER OF QUESTIONS RAISED

	Lecture (N = 36)		Self-Directed (N = 30)	
	Mean	S. D.	Mean	S. D.
Item 1 "What questions do you have about the field of psychology as applied to education? Put down as many questions as you can think of."	3.03	1.90	3.53	1.30
Item 2 "Read the following selection and then comment on it in at least 25 words." (Selection concerned problem of child in second grade.)	.53	1.29	.73	1.48
Item 3 "What problems or questions did the selection you just read (in Item 2) suggest to you?"	3.42	1.74	3.90	1.18
Item 4 "Is there anything in the field of Educational Psychology that you want to find out more about? What?"	.75	.80	1.03	.90
Item 7 "Read the following selection and then in the space below write as many questions as you can think of that the quotion suggests to you." (Quotation asserted need to emphasize mathematics and other "basic" subjects in the high-school curriculum.)	2.75	2.11	2.86	2.25
Sum of Questions Raised, Items 1, 2, 3, 4, and 7.	10.58	4.12	12.10	3.87
Merit of Questions (sum of judgments of merit of responses to Items 1, 2, 3, 4, and 7).	15.88	2.33	17.16	2.91

TABLE II

STUDENTS CLASSIFIED BY TOTAL NUMBER OF QUESTIONS RAISED AND BY TREATMENT

Total Number of Questions		SDS	Lecture
Low	(3 - 8)	4	12
Middle	(9 - 14)	17	19
High	(15 - 23)	9	5

Chi-square = 4.76
$.05 < p < .10$

Using the binomial expansion to calculate the probability of obtaining five out of five scores all in the predicted direction (7), the superiority of the SDS group is significant ($p < .03$). In addition, the SDS group was superior in the diversity of questions raised, in the total number of questions raised, and in the merit of the questions raised.

Table II shows the frequencies of individuals raising different numbers of questions, subdivided into low, middle, and high numbers of questions raised. Although the middle categories are of about the same size in both SDS and lecture groups, the proportion of individuals in the extreme categories is in accord with the hypothesis that the SDS condition stimulated greater curiosity as evidenced by question-raising behavior.

Interest—In four of the five items dealing with interest in the field of educational psychology, the SDS group was superior, and in four of the five items dealing with tendencies toward overt action indicating interest in educational psychology, the SDS group was superior. Thus, the SDS group was superior in 13 of the 15 items in the post-test questionnaire. Since the intercorrelations among scores in the various items were negligible, it is legitimate to treat these 15 scores as independent observations. Using the binomial expansion, this result is highly significant ($p < .004$).

Action-indicators—Within the group of "action-indicator" items, an interesting trend emerged. The SDS group reported a larger mean number of readings since the course than did the control group, while the control group reported a larger number of activities other than reading. This difference in pattern was highly significant (chi square = 10.95, $p < .001$). The readings reported included serious books, professional journal articles, articles in serious magazines (e.g., Harpers), and articles in slick magazines and newspapers. Other activities reported included observations of children in and out of school, observations of teachers and other people, application of course material to the teaching of piano and remedial reading, attending school board meetings, and attending lectures. Most of the activities other than reading were concerned with the application of what had been learned. Thus, the significantly different distribution of activities reported suggests that SDS students, in the year following the course, had tended to do more things to increase their knowledge, while the control students had tended more often to apply what they had learned.

Replication

Since the results of the experiment above were promisingly consistent, but of only tenuous statistical significance, it was decided to repeat the study with another course. The course chosen was the second semester of a two-semester sequence in General Psychology. As before, the SDS section met three times a week. The SDS section was organized in small discussion groups, usually of five or six students, which met at a place and time of their own choosing, with no instructor present.

Because the decision to repeat the previous study was taken after the course in General Psychology had already been given, arrangements could not be made to obtain comparable final examination scores from both groups. Instead, the groups were matched on the basis of the Academic Achievement Test (AAT) given to all freshmen and the results to be reported deal only with the post-test, given one year after the course was finished. The Ss in the post-test consist of 14 matched pairs; in every case either both members have identical scores on the AAT, or where identical scores could not be located, the pair was constructed so that the SDS member would have a slightly lower score, thus working against the hypothesis that the SDS group would exhibit superior performance in the post-test.

The post-test consisted of an objective test of retention of material covered in the second semester of General Psychology, items intended to elicit question-raising behavior, items dealing with interest in psychology, and items dealing with action-tendencies toward getting more information about psychology. Although all of these kinds of items have been described above, the specific items were modified to suit the changed material of the course. For example, in one of the question-raising items, the Ss were asked to comment on a passage describing the "authoritarian personality", rather than on a passage about educational psychology. Finally, the post-test was somewhat shorter than in the previous study, to reduce labor.

Results

On the retention test, the SDS group was slightly but insignificantly superior to the control group.

On each of three items dealing with question-raising behavior, the SDS group was superior in the number of questions raised. In the total number of questions raised, the superiority of the SDS group just failed of significance, using a two-tailed Wilcoxon test for paired replicates. The SDS mean for the total number of questions raised was 8.0; the control mean was 6.4.

On each of three items dealing with expressions of interest in psychology, the SDS group was superior to the control group. On one of these, the superiority of the SDS group was significant at the one percent level, using Wilcoxon's test for paired replicates. This item was a nine-point rating scale for the question, "In general, how interested are you in psychology?" The midpoint of the scale was five, the SDS mean equalled 5.41, and the control mean equalled 4.18.

There were three action-indicators, on two of which the SDS group was superior to the control group and one of which the reverse was true. None of these differences was significant.

Thus, on eight out of nine items, the SDS group was superior to the control group in interest and curiosity about general psychology.

Discussion

It seems clear that self-directed study produced a small but persistent superiority in question-raising behavior and in other indices of curiosity. In both studies when tested almost a year later, the SDS groups asked more questions and were prepared to exert greater efforts to seek information about educational psychology. Although they had only one-third the usual number of contact hours with the instructor, the SDS groups suffered no loss in retention of factual knowledge.

It is worth remarking that the instructors in the two SDS sections were quite different. There was a ten-year age difference between them. The older instructor would unquestionably be described as mature, steady, and well-organized—a man for whom the lecture might seem a natural medium. The younger instructor would be described as youthful, ebullient, and enthusiastic—a man who particularly enjoys contact with individual students. In spite of these differences, essentially the same results were obtained in both courses.

The question remains, when should we expect greater curiosity to lead to more enduring mastery of subject matter? Curiosity cannot lead to information-seeking and continued mastery of material if the environment presents the S with a multitude of stimuli that produce behavior patterns antagonistic to independent information-seeking. Thus, isolated experiments on self-directed study can have only limited effects. But if curiosity is not the only attribute of the permanently educated individual, it is at least a sine qua non. In this sense, curiosity should not be viewed as an end in itself but as an essential class of events in the stream of cognitive processes that characterizes the intellectually independent individual.

Surely a major goal of education is to stimulate the student to further pursuit of knowledge on his own initiative, after the compulsions of the classroom are far behind. Curiosity may be said to have a "gatekeeper" function in the educational system: if the system arouses further exploratory or reorganizing behaviors, it may set off a process which is self-sustaining and which may in large part determine the whole character and direction of the individual's future life.

Summary

Two classes in educational psychology, equated with respect to aptitude, were used in the present experiment. One class was taught by the traditional lecture method, three lectures per week, while the other was handled by a self-directed technique, involving a considerable amount of small-group discussion and only one lecture a week.

There was a small but non-significant difference in favor of the self-directed group in the mastery of course material. Ten months after the course was completed the Ss were reassembled for post-tests covering course material and measures of curiosity. The self-directed Ss were still slightly superior in retention of course material. On 13 of 15 items measuring curiosity, the self-directed Ss were superior. The study was repeated in a somewhat briefer form in a course in general psychology with substantially the same results.

The technique employed in the present experiment of placing a major responsibility on the student for his own education suggests interesting possibilities for developing attitudes toward learning which will result in the student's continuing a search for knowledge after the formal classroom experience is over.

REFERENCES

1. Berlyne, D. E. "An Experimental Study of Human Curiosity," British Journal of Psychology, XLV (1954), pp. 256-65.
2. Bloom, B. S. "Thought Processes in Lectures and Discussions," Journal of General Education, VII (1953), pp. 160-69.
3. Hovey, D. E. The Development of Curiosity Under an Experimental Method of Teaching the Educational Psychology Course, unpublished MA thesis, University of Colorado, 1958.
4. Kersh, B. Y. "The Adequacy of 'Meaning' as an Explanation for the Superiority of Learning by Independent Discovery," Journal of Educational Psychology, XLIX (1958), pp. 282-92.
5. Liverant, S. "The Use of Rotter's Social Learning Theory in Developing a Personality Inventory," Psychological Monographs, LXXII (1958), pp. 1-23.
6. Stroud, J. B. "Experiments on Learning in School Situations," Psychological Bulletin, XXXVII (1940), pp. 777-807.
7. Walker, Helen, and Lev, J. Statistical Inference (New York: Henry Holt Co., 1953).

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 7, March 1963)

Student Perceptions of Group Work in Teacher Education Classes

EDMUND AMIDON and ARTHUR BLUMBERG
Temple University

THOUGH RELIABLE evidence is not available, one gets the impression that, in recent years, the use of one or more "group techniques" in teacher education classes has increased markedly. For example, one instructor in a nationally known graduate college of education estimated that the use of buzz groups or discussion groups in course work in that institution had increased by approximately 60 percent in the space of about five years. Indeed, one is hard put to encounter students on either a graduate or undergraduate level who have not had rather wide experience in the classroom use of groups or group techniques, varying from large class discussion to buzz groups, project groups, role-playing, panels, etc.

Such an apparent widespread increase in the use of group teaching methods prompts one to assume that weighty evidence is at hand to suggest that "groups are good." An examination of the research, however, reveals that this is not necessarily the case at all. What we are confronted with is a situation of conflicting, inconclusive research results. Bovard has summarized the problem in the following way:

> Available experimental evidence indicates that the amount of student-to-student interaction in the classroom has no statistically discriminable effect on a number of variables, including learning of objective content, liking for the classroom teacher, and evaluation of the teacher's performance by the students. The same evidence, however, clearly indicates that the amount of such interaction has a decided effect on such non-content variables as perception of an objective stimulus, liking for other students, and spontaneity of behavior.[1*]

As things stand, then, it would appear that as far as that which lends itself readily to measurement, i. e., the amount of objective subject matter learned, there appears to be no particular justification to depart from traditional teaching methods. On the other hand, when consideration is given to the intangible, hard-to-measure effects of classroom interaction, the situation is not nearly as clearcut.

Studies have also been made that have attempted to appraise student reactions to the lecture method of college instruction as contrasted to group methods. Again, the results are in conflict. Wispe[2] for example, found that a majority of the students he studied preferred directive teaching with an examination emphasis. Contrariwise, Brooks and Davis reported, "College students appear to be ready for more varied and differentiated teaching procedures than are probably being used in many college classrooms."[3]

The aim of the present study is based on the acceptance, on a nonvalue basis, of the assumption that a great deal of group methodology is current in teacher education instruction. With this point in mind, the writers have attempted to get an appraisal, from the student's point of view, of different kinds of group teaching methods or techniques which they have experienced in their regular curriculum. Specifically, interest focused in the following areas:

1. Differences between elementary and secondary education students in regard to their perception of effectiveness of different methods.
2. The factors seen by these students as having relevance for making a particular method successful or unsuccessful.
3. Level of satisfaction with different methods.
4. The extent to which these students see themselves using group methods once they are in their own classroom as teachers.

While the writers recognize the effect of certain variables—the teacher's personality, for example—on the learning process, it was felt, nevertheless, that the kind of information that might result from this study would help shed more light on the dynamics of classroom interaction.

All footnotes will be found at end of article.

Methodology

Data for the study was collected by means of a questionnaire that was administered to sophomores and juniors enrolled in both the elementary and secondary education curriculum at the College of Education, Temple University, during the spring semester, 1960. These particular students were registered in a required course in group processes. The total N was 114. Of this number, 56 were preparing for elementary school work and 58 for secondary.

The questionnaire was first constructed in a more or less open-ended fashion and administered to one section of students. They were asked to rate six particular group techniques for effectiveness as they saw it and then give their reasons why they picked their first choice and why their last. From this initial test run the "technique choices" were condensed from six to four, and rating scales were constructed that focussed on those factors of a dynamic nature that were perceived by students as being related to the effectiveness of a technique and their satisfaction with it. The final four techniques chosen for analysis were: large class discussion with the teacher in control, small buzz groups, student project groups, and role-playing. The dependent factors to be studied were: opportunity for students to make verbal contributions, locus of classroom control, the amount of classroom security and control, accomplishment of "solid" subject matter learning, opportunity to clarify one's ideas, and opportunity for a mutual sharing of ideas with other students.

In addition, questions were inserted concerning the student's overall satisfaction with group techniques in teaching, their rating of a straight lecture as compared to their first group technique choice, and their perception of the extent to which they saw themselves using group techniques when they became teachers.

Analysis

A. Initially, for the purposes of analysis, the students tested in this study were considered in two different classifications. The first classification was concerned with whether or not the students were enrolled in either the secondary or elementary education curriculum. The second classification was based on the student's feeling about small group work in the classroom. That is, they were separated by whether they preferred general classroom discussion or some "small group" teaching method.

In this first analysis there were four categories: 1) elementary students who preferred general class discussion, 2) elementary students who preferred small group methods, 3) secondary students who preferred general class discussion, and 4) secondary students who preferred small group methods.

Using a two-by-two contingency table, a chi-square was computed to find out if there were significant differences among frequencies presented in the various cells (Table I).

The value found in this test of significance indicates that if we accept the .05 level of confidence, the observed distribution appears not to be a chance one. It would seem likely that there is reason, other than chance, why so many more elementary students prefer the small group methods and, conversely, why so many more secondary students prefer the general class discussion.

B. The second major analysis is concerned with trying to determine why students prefer one method rather than another.

As we have noted, before the questionnaire used in this study was administered, an open-ended questionnaire was administered to a group of students. The purpose of this questionnaire was to find out what significant characteristics made a student like or dislike group teaching methods. On the basis of a content analysis the following six characteristics appeared to be crucial:

1. Opportunity or lack of opportunity to participate verbally
2. Opportunity or lack of opportunity to get ideas clarified.
3. Opportunity or lack of opportunity to share ideas
4. Amount of security felt
5. Locus of classroom control
6. Amount of "solid" subject matter learning

Each of these six issues was built into a nine-point scale, and students were asked to react to each scale depending on how they felt about the issue it presented. An example of the scale used is shown at the top of the next page (considering the type of group teaching methods preferred). Similar scales were built for the other five items.

The major analysis of this part of the questionnaire was to compare the responses on these scales of those students who liked large group discussions with those who preferred small group methods. Means and results of the t-test run on the various scales are presented in Table II.

These analyses will, in general, answer the question as to which of the six factors mentioned above were important in determining which method the students liked better—small group methods or large class discussion.

1. The first item dealt with verbal participation. The means of the two groups are significantly different from one another if we accept the .01 level of confidence. The people who prefer the small group methods seem to feel they get a better chance to participate verbally in this situation than do the people in the large group discussion.

2. As far as the locus of control in the classroom was concerned, there was also a significant difference, .01 level, between the responses of those who preferred the large group and those who

TABLE I

GROUP METHOD PREFERENCES OF ELEMENTARY AND SECONDARY EDUCATION STUDENTS

	Number of Students Preferring General Class Discussion		Number of Students Preferring Group Methods		
	Observed Frequency	Theoretically Expected Frequency	Observed Frequency	Theoretically Expected Frequency	Total
Secondary Students	40	33.6	18	24.4	58
Elementary Students	25	32.4	31	24.6	56
Total	65		49		114

$X^2_{obs} = 5.74 < .02$

TABLE II

FACTORS RELATED TO STUDENT PREFERENCE OF GROUP TEACHING METHODS

Question	Means: All Students Choosing Large Group Methods	Means: All Students Choosing Small Group Methods	t	Probability
Verbal Contribution	8.20	7.01	2.33	.01
Maintaining Control	7.37	2.99	7.10	.01
Security and Comfort	8.40	6.40	4.70	.01
Subject Matter	3.53	2.97	1.30	...*
Clarifying Ideas	7.90	7.30	1.2	...*
Sharing Ideas	2.50	3.40	2.1	.01
Satisfaction	6.94	6.29	1.5	...*
Group vs. Lecture	4.22	4.12	.27	...*
Group Work in Own Teaching	4.31	3.29	3.1	.01
Sharing Ideas	3.27	4.57	5.00	.01
Satisfaction	7.53	5.47	4.20	.01
Group vs. Lecture	3.30	4.77	2.40	.01
Group Work in Own Teaching	4.07	3.27	3.50	.01
Sharing Ideas	2.93	2.05	2.40	.01
Group Work in Own Teaching	4.53	3.94	3.10	.01

*Probability level less than .05.

As far as maintaining control in the classroom is concerned—

the teacher is almost always in control	sometimes the teacher is in control, sometimes not	the teacher is rarely in control

chose the small group. The large group choosers said that they liked it because the teacher exercises more control. Those who preferred small group methods indicated that they were influenced in this choice by the fact that the teacher did not exercise as much control as in the large situation.

3. There was a significant difference on the security and comfort item between the large group people and the small group people. Small group preferrers felt more comfortable in this situation than did the large group people in the one they liked best.

4. The fourth item was concerned with "solid subject matter learning." This apparently was not a crucial point as the two groups answered it in ways not significantly different from one another.

5. The fifth issue dealt with clarifying ideas. Again there was no significant difference in the way the two groups responded.

6. The sixth scale focussed on opportunity for sharing ideas. There was a significant difference, .01 level, between the large and small group choosers in the way they responded. Small group people felt there was a greater opportunity to share ideas in that situation than the large group people felt there was in the general class discussion.

If we compare those secondary education students preferring large group methods with those elementary education students preferring large group methods, we also find a difference, .01 level, on this same item, i. e., the opportunity to share ideas. The large group elementary feel there is less opportunity to share ideas in the large group than do the large group secondary people.

If we compare small group elementary with small group secondary, we find the elementary seeing more opportunity in the small group for sharing ideas than do the secondary.

C. A third section of our analysis was that part of the questionnaire concerned with general level of satisfaction with group methods.

1. The first question was concerned with general satisfaction with group work in college teaching. If we compare those who preferred the large with those who liked the small group methods, there is no significant difference in regard to general level of satisfaction. If, on the other hand, we compare the elementary students who preferred the large group against the secondary with a similar choice, the elementary students appeared to be more satisfied. In addition, a comparison of all the elementary students with all the secondary students indicates that the former are more satisfied.

2. The second item in the "satisfaction" area was concerned with straight lecture as compared with group methodology. Comparing those who chose the large group situation with those who chose the small group yields no significant difference. However, if we look at elementary students who preferred the large group and compare them with their counterparts in the secondary group, a difference, .01 level, is found. The secondary people feel that the lecture, in comparison with group methods, is more effective than do the elementary people.

3. Finally, the students were queried as to the extent they saw themselves using group methods in their own classroom when they assumed teaching positions. Here there are significant differences between the large and small group preferrers. The results, .01 level, seemed to indicate that the latter would be more likely to use group methods. In regard to the choice of elementary and secondary people, the results indicated that the elementary students, whether they chose the small group or the large group situation, felt they would use group methods more than the secondary students, .01 level.

Interpretations and Conclusions

It would appear from the results of the study that the students in elementary education are more apt to prefer small group methodology in their classrooms than are students in secondary education. As can be noted, this thread runs through all the statistically significant results in the previous section. About 50 percent of the elementary students responding to the questionnaire indicated that they preferred small group work in their classrooms, whereas approximately 30 percent of those students in secondary education indicated their preference for small group work (chi square significant at .02 level).

There are probably many reasons why this situation exists. For example, the secondary student is probably more apt to see himself as a conveyor of a particular branch of subject matter than is the elementary education student who is more likely to see himself as working with children in a school setting. The traditional manner of conveying subject matter is through the use of a lecture and recitation whereas commonly accepted classroom methods in the elementary school usually have an emphasis on some kind of group learning situation. It is possible that the different emphases put on the methods of instruction in the public schools are reflected in the kinds of instruction that are used in the college classroom by both secondary and elementary education instructors. The point here is that it is rather common knowledge that most teachers teach as they in turn have been taught. In addition, it is possible that although use is made of group methods in secondary education instruction in college, there are more success experiences in this regard in elementary education because the instructors themselves are more committed to this method of instruction both in college and in the public schools.

It becomes necessary, however, to go beyond generalized speculation concerning the problem and analyze some of the specific dynamic variables that were perceived by the students in our study as being directly related to their choice of one method of classroom instruction as opposed to another.

Verbal Participation—Our comparison of students who preferred small group methods over large group methods indicated that an important consideration was that the former gave opportunity for more to participate verbally than did the latter. This would tend to indicate that people who prefer small group work see verbal participation as being important to the learning process. In addition, it is possible that these people see verbal participation as a means of getting more people involved in the classroom experience and that they see this kind of involvement as being crucial to maximum learning.

Maintaining Control—The second significant variable concerned with student preference of small group as opposed to large group methods is related to where the students like to see the locus of control in the classroom. Those people who prefer large group methods apparently wish the instructor to be in control at all times. The students who preferred the small group methods in the classroom seem to indicate that there should be situations during the course of instruction when students are responsible for their own control of classroom processes. It is possible to interpret these results as a desire on the part of some students to rebel against authority, that is the authority of the teacher's control. However, it seems much more likely that this desire on the part of some students to exercise more control in the classroom through the use of small group methods is related to the need to be independently creative within the classroom context and share in the working out of their own destinies. The large difference between the means on this scale would seem to indicate that the issue of control in the classroom is seen by students as being perhaps the most crucial one in their choice of particular kinds of classroom methods. It would not be out of bounds to speculate that preferences in this regard will be carried over into the actual teaching situation when these students are employed in the public schools and that their concerns for control will be reflected in the manner in which they conduct their own classroom.

Security and Comfort—An interesting difference in our findings concerned the point that those people who preferred the small group setting in classroom instruction saw more security and comfort in this setting than did the large group preferrers see in their choice. The issue seems to be that those who chose small groups are more concerned about security and comfort in the learning situation than are those who chose large groups. In a sense, we are saying that those people who prefer small groups have made a more realistic and sensitive assessment of primary needs within a classroom situation than did the others. Perhaps these findings indicate that the students who preferred the small group method are actually more concerned with meeting the needs for security and comfort than the others and that they see small group methodology in the classroom as a means of doing it.

Sharing Ideas—The people who chose the small group setting saw it as more of a vehicle within which they could share their ideas with each other than did those who chose the large group setting. This result appears to be consistent with the rest of the findings of the study. Particularly related to this factor would be the one cited in regard to the need for verbal participation. That is, apparently it is seen by these students as essential to the learning process that not only must they be able to get ideas and information from the instructor but that they also be able to share their own ideas with others and in turn get ideas from their peers.

In connection with this same variable, when we compare elementary education students who preferred the large group situation with secondary education students with the same preference we find that the elementary students thought there was more opportunity for sharing ideas in this situation than did the secondary students. This would seem to reinforce our previous conclusions concerning the needs for elementary education students to have a sharing of ideas take place within the classroom setting. Likewise, if we compare both elementary and secondary students who chose the small group situation, we find the same set of conditions existing regarding the need to share ideas in the classroom. That is, the elementary students seemed to be more concerned that this kind of situation exist than did the secondary students.

Satisfaction with Group Teaching Methods—If we compare all the elementary students with all the secondary students, we find that there is a significant difference in their satisfaction with group teaching methods, the elementary students being more satisfied than those in secondary education. This, too, will reinforce our previous findings, and it lends credence to our earlier hypothesis concerning the point that elementary education students would be more likely to choose group methods as a way of teaching than would secondary education students.

Comparing Straight Lecture Method with Group Teaching Methods—As might have been predicted, the elementary education students preferred the group teaching method as compared with the straight lecture relatively more than did the secondary education students. In substantiating their feelings by means of an open-ended question that followed the questionnaire item, they gave reasons that seemed to reinforce our previous findings. For example, reasons give for liking group methods in preference to the straight lecture were concerned with the feelings that in the use of groups one could achieve an exchange of ideas, one could become more involved, there was more freedom of expression, and one had more of a feeling of independence. By the same token, their feelings about a straight lecture were that there was no opportunity for an exchange of ideas, there was not as much security in the classroom, there was no chance for free expression, and that, frequently, straight lectures tended to be boring. Reasons given by secondary education students for their choice of straight lecture method were that they felt that they could get the facts better in a lecture, they felt there was more control in the lecture situation, they felt that the lecture was more systematic, and the feeling existed that they were able to get ideas from the expert.

Perceptions as to the Future Use of Group Methodology in the School Classroom—As might be expected, there was a significant difference between people who chose the small group and those who chose the large group as a most preferred method of teaching in regard to the extent to which they saw themselves using group methods when they are on the job. The people who preferred the small group methods indicated that they would be more inclined to use them frequently in their own teaching than did those who chose the large group. This finding would tend to lend weight to our earlier point that teachers probably lean toward teaching in the way that they have been taught. However, we would add to this the idea that teachers probably tend to want to teach in a way that they found satisfying when they, in turn, were students.

On this same item, the differences that occurred between the elementary students as a whole and the secondary students as a whole would further reinforce this idea.

Conclusions

Based on the findings of this study, it would appear to be possible to make the following conclusions:

1. Apparently, one's curriculum makes a difference in the way one perceives various kinds of group methodology in teaching.
2. It appears that students who prefer small group methods as compared to those who like large group procedures differ in the degree to which they see certain factors as being crucial to their choice. The factors are a) opportunity to participate verbally, b) locus of classroom control, c) amount of security and comfort, and d) opportunity for sharing ideas.
3. There was also a difference between the two groups in regard to a) satisfaction with group methods, b) perceptions of value of group methods as compared with a straight lecture, and c) anticipated use of groups methods on the job.

Two final points seem in order:

1. Further research is needed in order that some of the questions raised in this study may be investigated in more depth.
2. There seems to be sufficient evidence in the present study to indicate that college instructors need to rethink the aims of their teaching in regard to the methodology they use in order to accomplish these aims.

FOOTNOTES

1. E. W. Bovard, Jr. "The Psychology of Classroom Interaction," Journal of Educational Research, XLV (November 1951), p. 215.
2. L. G. Wispe, "Evaluating Section Teaching Methods in the Introductory Course," Journal of Educational Research, XLV (November 1951), pp. 161-86.
3. L. A. Brooks and L. R. Davis. "Student Opinion Regarding Instructional Procedures on the College Level," Teachers College Record, LXI (March 1955), p. 337.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 7, March 1963)

Lola Gets What Lola Wants From Supervision

CLIFTON M. CLAYE
Texas Southern University

THE EXTENT to which supervisors provide those services desired by teachers has been investigated by several researchers. Rosenthal (7), Replogle (6), Meil (5), and Kyte (4) have been rather specific in defining the job of the supervisor in terms of what she does as well as the services desired by teachers.

This study was undertaken not only to find out the kinds of services desired by teachers and the extent to which such services are received, but also to determine the nature of the content of courses in supervision at Texas Southern University. The questionnaire was designed to test the following hypotheses:

1. There is no significant difference in the frequency of desire and receipt of the supervisory services as represented by the items in the questionnaire.
2. There is no significant difference in the proportion of subjects desiring the supervisory services represented by the items in the questionnaire with reference to the title of the person used as a frame of reference by them.
3. There is a significant positive relationship between the proportion of subjects desiring the supervisory responsibilities represented by the items in the questionnaire and the length of time the subject has been in service.

The items making up the questionnaire were divided into six categories: 1) Teacher Growth Activities, 2) Teacher Participation Activities, 3) Supervisor's Professional Competency, 4) Supervisor's Activities, 5) Supervisor's Personal Qualities, and 6) Human Relations Activities.

In the investigation of the desire and receipt of the supervisory responsibilities represented by the items in the questionnaire in each category, the terms employed are defined as follows:

A. Supervisor — The person in any school or school district whose major responsibility is the improvement of instruction in that school or school district.
B. Supervision — The effort to stimulate, coordinate, and guide the continuous growth of teachers in a school or school district so as to improve the quality of instruction therein.
C. Lola — A classroom teacher. (Hypothetical)
D. Teacher Growth Activities — Those activities designed to insure the professional growth of teachers on the job.
E. Teacher Participation Activities — Those activities in which teachers and supervisors work cooperatively to improve the teaching-learning situation.
F. Human Relations Activities — Those activities designed to insure improved inter-group relationships among the personnel in a given school or school district.
G. Supervisor's Personal Qualities — Those qualities or traits which identify the supervisor as a person.
H. Supervisor's Professional Competency — Those activities on the part of the supervisor which indicate his ability to carry out his function.
I. Supervisor's Activities — Those activities which describe what the supervisor does.
J. Frequently — Help desired or received as represented by a particular item at least three-fourths of the time.
K. Occasionally — Help desired or received as represented by a particular item approximately one-half of the time or less.
L. Not At All — Help desired or received as represented by a particular item one-fourth of the time or less.

Procedures

The study involved a survey of 79 teachers enrolled in the graduate school at Texas Southern University during the first term of the 1959 summer session. Employees of the public schools in five different states were represented in the sample, and they represented teachers of all grade levels in the elementary school and nine different subject-matter areas in the secondary school.

The length of time the subjects had been in service was divided into four groups. Thirty-seven and three-tenths percent of the subjects had been in service zero to three years; 29 and one-tenth percent,four to six years; eleven and four-tenths per cent, seven to ten years; and 22 and two-tenths percent had been in service eleven or more years.

The enrollment of the schools represented by the subjects ranged from less than 100 (three schools) to more than 2000 (ten schools). When the sizes of the schools represented by the subjects were compared in terms of the number of teachers employed, the range was from two one-teacher schools to five schools with 95 or more teachers.

Our definition of a supervisor did not limit us to any given title to be used as a frame of reference by the subjects. Nineteen percent of them used the "special supervisor" as a frame of reference, 28 percent used the "district supervisor," 45 and six-tenths percent used the "principal," and seven and four-tenths percent used the "dean" and "department head."

It was reasoned that the desires of the subjects because of their various backgrounds and experiences would represent the desires of other teachers relative to supervision.

At the time of administering the questionnaire, subjects were told that their participation was voluntary and that the survey had nothing to do with their schooling and in no way would affect their grades. No names were required. However, subjects were asked to reveal certain information relative to the schools in which they worked, the length of time they had been teaching, and the title of the person used as a frame of reference. *

Findings and Discussion

Since the hypotheses dealt with the differences in the proportion of subjects desiring and receiving the supervisory responsibilities represented by the items, the test of the hypotheses involved the use of chi-square. The five percent level of confidence was used.

A comparison of the percent of subjects desiring and receiving the listed supervisory responsibilities for all the subjects regardless of the title of the person used as a frame of reference that: 1) 56 percent of the subjects desired the services frequently while 30 percent of the subjects received them frequently; 2) 36 percent of the subjects desired the services occasionally while 43 percent received them occasionally, and 3) eight percent of the subjects did not desire the services at all while 27 percent did not receive them at all.

It was postulated that there was no significant difference between what teachers want from supervisors and what they get from supervisors.

Chi-square was significant for the differences in receipt and desire of the listed supervisory responsibilities for ten degrees of freedom. However, when the differences in the proportion of subjects desiring and receiving the services were compared on the basis of the title of the person used as a frame of reference, the differences found for the special and district supervisors were not significant, but were in the case of the principal. number of subjects using other titles as a frame of reference was too small to test.)

Of those subjects using the principal as a frame of reference, 62 percent desired the items frequently, 33 percent desired them occasionally, and five percent did not desire them at all. At the same time, 28 percent received them frequently, 44 percent received them occasionally, and 28 percent never received them at all.

Of those using the special supervisor as a frame of reference, 53 percent desired the items frequently while 44 percent received them frequently; 32 percent desired the items occasionally while 35 percent received them occasionally, and 15 percent did not desire them at all while 21 percent never received them at all.

Of those using the district supervisor as a frame of reference, 49 percent desired the items frequently, while 38 percent received them frequently, 43 percent desired the items occasionally, while 48 percent received them occasionally, and eight percent did not desire the items at all while 14 percent did not receive them at all.

The significant differences in the desire for and receipt of the supervisory responsibilities as they relate to the principal might indicate one or more of the following to be operative:

1. Principals are not aware of the available knowledge relative to supervision.
2. Principals are aware of this knowledge but prefer a comfortable and comforting set of routines.
3. What principals do is so circumscribed by imposition of rules and regulations from the central office that the extent of their supervisory activities is limited.
4. Principals do not know what to do.
5. Principals do not know how to get started.
6. Principals feel that they know what is good and desirable for teachers, hence, what they do or do not do is the best thing for the teacher.

---

* The Questionnaire used is not presented in this report. However, copies may be obtained from the author.

7. Principals' concept of supervision ignores all we know about the social process, hence about supervision.
8. Principals believe that their major responsibility is one of administering — not improvement of the teaching-learning situation. Hence, improvement of the teaching-learning situation is secondary to keeping teachers and learners under control.

Further analysis of the differences in desire and receipt of the supervisory responsibilities as they relate to the principal reveal chi-square to be significant in each category. These findings support the conclusions of Briggs (2), Sheelam (8), and Hansen (3). These findings might also indicate that teachers and principals are not together on what each expects of the other. Adams and Dickey (1) suggest that teachers and principals need to understand each other clearly. They call this process the "exercising of leadership," and they suggest that this leadership base should be broadened by utilizing the full potential of teachers.

Further analysis of these differences as they relate to special and district supervisors shows that chi-square was significant for "Category B"– Teacher Participation Activities.

The insignificant values of chi-square as they relate to the special and district supervisors might suggest that since persons with these titles are released from many of the details associated with running a school, they are in a position to devote more time and effort in helping teachers. It might also suggest that persons having these titles have had more specialized training in the process of supervision, hence are in a better position than principals really to help teachers. Too, it might suggest that persons using these titles as a frame of reference understand the function of the supervisor better than those using the principal.

This hypothesis was rejected when desire for and receipt of the services represented by the items was considered for the entire group. However, when the differences were considered in terms of the title of the person used as a frame of reference, the hypothesis was accepted for special and district supervisors and rejected for principals.

It was postulated that there was no significant difference in the proportion of subjects desiring the listed supervisory services in regard to the title of the person used as a frame of reference. When the differences in the proportion of subjects desiring the supervisory responsibilities represented by the items regarding the frequency of desire and the title of the supervisor used as a frame of reference were tested — none was found to be significant. This fact might suggest the need for one or more of the following:

1. The same type of training, understanding, and insight as to methods and techniques of improving the teaching-learning situation for principals, special supervisors, and district supervisors.
2. The hiring of more clerks in order that principals might be relieved of some of their routine duties so that they might devote more time to supervision.
3. Principals take a "look-see" at what they are doing and the amount of time they spend doing it in order to eliminate routines which do not contribute to the improvement of the teaching-learning situation.
4. Re-orientation on the part of principals as to their particular role and the desirable behavior patterns in administering schools.
5. Recognition by principals that they can no longer carry the "big stick" and get the job done.

The proportion of subjects desiring the listed supervisory services from the title of the person used as a frame of reference was further analyzed and tested for significance. None was found to be significant. This hypothesis was accepted.

It was postulated that there is a significant positive relationship between the proportion of subjects desiring the supervisory responsibilities represented by the items and the length of time subjects had been in service. The differences in frequency of desire by the four groups were compared. None of them was found to be significant.

These findings should suggest that those of us who are prone to say that older teachers are not concerned with the improvement of the teaching-learning situation are mistaken. All teachers need and want supervision. This hypothesis was accepted.

Summary

Effective supervision is based on sound principles of social change and group dynamics. To effect change in any social process, knowledge of the desires of the individuals involved in the process is important. Unless the individual desires change, no permanent change will be effected. These findings, on the basis of the population studied, indicate that teachers want effective change in the teaching-learning situation. This is the ultimate end of supervision. This study was based on the belief that teachers want to improve the teaching-learning situation and that supervisors have a responsibility for helping in the process. This help, we believe, can be given through the carrying out of the supervisory responsibilities represented by items used in this study.

The sample of subjects used in this study was too small to draw any valid conclusions. However, the findings do suggest that:

1. Teachers want supervision from principals as well as from those persons with titles of supervisor.
2. Principals do not supervise adequately.
3. The kinds of help teachers want do not change significantly as the length of time in service varies.
4. The kinds of experiences provided for principals in our training programs need reexamination.

REFERENCES

1. Adams, H. P., and Dickey, F. C. Basic Principles of Supervision (New York: American Book Co., 1953).
2. Briggs, Thomas H. "A Self-Rating Scale for School Principals," Bulletin, National Association of Secondary School Principals, XXVII (December 1943), pp. 49-64.
3. Hansen, Basil C. "A Teacher Looks at Her Supervisor," School Executive, LXVI (May 1947), p. 42.
4. Kyte, George C. "This is the Kind of Supervision Teachers Want and Appreciate," Nation's Schools, XLVIII (July 1951), pp. 33-34.
5. Meil, Alice. "Supervisors I Have Known," Educational Leadership, X (November 1952), p. 95.
6. Replogle, V. L. "What Help do Teachers Want?", Educational Leadership, VII (April 1950), pp. 445-449.
7. Rosenthal, S. S. "An Open Letter to My Supervisor," The Instructor, LXI (September 1951), p. 11.
8. Sheelam, Mary A. "How Can the Supervisory Responsibilities of the Principal Become Maximally Effective?" Bulletin, National Association of Secondary School Principals, XLI (April 1954), pp. 91-94.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 7, March 1963)

Relationship of Selected Adjustment Factors, College Ability, and Achievement to Drop-Outs and Nondrop-Outs Of College Freshmen

WARREN J. VORREYER
Southern State Teachers College
Springfield, South Dakota

COLLEGE drop-outs represent a tremendous waste of our most able and competent manpower. In addition to the significant waste of manpower, the cost of the drop-out student has placed undue burden on many institutions of higher education. Numerous reasons are given by the college student for withdrawal (1:18). It should be noted, however, that college drop-outs usually do not occur purely from one isolated element or condition. College student mortality is a most complex situation, involving several highly significant and interacting factors.

The present study was undertaken to investigate the relationship of home adjustment, health adjustment, social adjustment, emotional adjustment, scholastic achievement, and college ability to student retention and withdrawal, and to compute, if possible, a series of equations which could aid in the identification of potential drop-outs. Coefficients of correlation were also computed to determine the relationship of the first semester grade average to a) selected adjustment factors of the Bell Adjustment Inventory, b) college ability as measured by the Ohio State University Psychological Examination, and c) the high-school cumulative grade average of college freshman students.

It is hoped that the results of this multi-directional study may be of value to high-school and college counselors and guidance workers, to college admissions committees, and to the institution itself as a means of assessing its strengths and weaknesses. For some colleges, it might offer a practical means of controlling admissions to a suitable enrollment level.

Methods and Results

During the freshman orientation week of a Rocky Mountain University, more than 900 freshman students were administered the Bell Adjustment Inventory. From the office of student personnel, individual cumulative grade averages and ability scores were obtained. Available high-school cumulative grade averages were secured for each individual from the office of the registrar. The total number of freshman students for whom all scores were available was 732. After the sample had been established, it was divided into two groups, male and female. Each group was subdivided into enrollment periods A and B—Group A, first semester; Group B, first and second semester. Further, the groups were classified into subsections for each enrollment period; nondrop-out female, drop-out female, nondrop-out male, and drop-out male. The six independent variables selected for use were the four adjustment measures of the Bell Adjustment Inventory, O. S. U. Psychological Examination raw score, and the high-school cumulative grade average. The criterion was the first-semester university grade average.

From the computation of all intercorrelations, it was found that only two predictive variables had correlations of .53 or greater with the criterion. These variables were the high-school grade average and the O. S. U. Psychological Examination. The correlations were .61 and .53 respectively. The data are presented in Table I. Correlations of .09510 or above are significantly different from zero at the .01 level of confidence. Only the high-school grade average and the O. S. U. raw score differed significantly from zero at the .01 level. Thus, the following regression equation was derived:

$$Y = 1.2756 + .8467X_1 - .0035X_2$$

Y = First semester university grade average
X_1 = High-school cumulative grade average
X_2 = (raw score) O. S. U. Psychological Examination

Two series of equations of discriminant analy-

TABLE I

SIMPLE CORRELATIONS AMONG ALL VARIABLES

(N = 732)

		1 High-School Grade Average	2 Ohio Total Raw Score	3 Home Adjustment	4 Emotional Adjustment	5 Social Adjustment	6 Health Adjustment	7 First Semester Grade Average
High-School Grade Average	1.	1.000	.5086*	.0693	.0359	.0892	.0022	.6130
Ohio Total Raw Score	2.		1.0000	.0466	.0035	.1289	.0045	.5262
Home Adjustment	3.			1.0000	.3770	.1636	.3512	.0452
Emotional Adjustment	4.				1.0000	.4270	.5421	.0193
Social Adjustment	5.					1.0000	.2101	.0591
Health Adjustment	6.						1.0000	.0726
First Semester Grade Average	7.							1.0000

*Underlined correlations are negative.

TABLE II

EQUATION OF DISCRIMINANT ANALYSIS
(Female)

$$-.00380622x_1 + .00011155x_2 - 00011365x_3 - .00008907x_4 + .00000950x_5 - .00015602x_6 = \text{derived score}$$

If the predicted derived score is equal or greater than -.00203672, then the student should be placed in the group expected to drop out of college. If the predicted derived score is less than -.00203872, then the student should be placed in the group expected to remain in college.

x_1 = High-School Grade Point Average
x_2 = Ohio State University Psychological Exam - raw score
x_3 = Home Adjustment, Bell Adjustment Inventory
x_4 = Emotional Adjustment, Bell Adjustment Inventory
x_5 = Social Adjustment, Bell Adjustment Inventory
x_6 = Health Adjustment, Bell Adjustment Inventory

TABLE III

RESULTS OF T^2-TEST USED TO DETERMINE WHICH VARIABLES WERE RESPONSIBLE FOR THE SIGNIFICANCE OF THE DISCRIMINANT EQUATION
(Male—Second Semester)

Variable	T^2-Value
1. High-School Grade Average	47.9414
2. Ohio State University Test Raw Score	46.8800
3. Home Adjustment, Bell Adjustment Inventory	2.8404
4. Emotional Adjustment, Bell Adjustment Inventory	.2054
5. Social Adjustment, Bell Adjustment Inventory	4.0811
6. Health Adjustment, Bell Adjustment Inventory	.1789

At .01 level of confidence T^2 must be 6.69 or greater with $n_1 + n_2 - 2$ degrees of freedom.
At .05 level of confidence T^2 must be 3.86 or greater with $n_1 + n_2 - 2$ degrees of freedom.

TABLE IV

PROBABILITIES OF SURVIVAL STATUS FOR FEMALE STUDENTS FOR WHOM DERIVED SCORES ARE PREDICTED

Derived Scores Female	Probability of Remaining in College
-.000012728	99 in 100
-.00063553	90 in 100
-.00092443	80 in 100
-.00135565	70 in 100
-.00169432	60 in 100
-.00203872	50 in 100
-.00350448	40 in 100
-.00527461	30 in 100
-.00684807	20 in 100
-.00831601	10 in 100
-.01086462	1 in 100

sis were computed in order to determine whether drop-out students differed significantly from other freshmen enrolled. The first series considered only the first-semester sample data, in which a male equation and a female equation were derived. Each sex group was divided into drop-out and non-drop-out categories. Student survival status was determined empirically. The second series was identical in character, except that this series considered the total first and second semester enrollment data. All the forementioned variables were utilized in this calculation.

"T^2-tests" were applied to all equations of discriminant analysis in order to determine which variable or groups of variables contained the greatest amount of discriminatory power. However, it should be remembered that just because an isolated variable is not significant to the given equation at a specific level of confidence does not mean that its contribution to the function is without value. In many instances, the constellation pattern formed by the combination of all variables has greater value in prediction than the limited combination of highly significant variables.

Predictive Validations

A number of predictive validations were attempted in an effort to attain the more successful method. In addition to the forementioned functions, a similar regression equation (2:45) utilizing the additional variable derived from the Iowa High School Content Examination—social studies raw score—was employed. It was found that when the first semester discriminant femal equation was applied to the female drop-out group, it predicted 80 percent accurately. The additional regression equation applied to the first semester male drop-out group predicted 83 percent accurately. Predictive validations were computed for all the equations. Finally, considering all validations, the first semester female discriminat equation for female students and the regression equation utilizing the social studies raw score for male students were the best predictive instruments. Tables displaying the probabilities of survival status for male and female students in reference to their respective equations were constructed.

Conclusions

Subject to the limitations of the study, the following conclusions were drawn: 1) Students who remain in college until graduation or students who terminate their programs before graduation tend to possess certain clusters of characteristics which enable these groups to be identified. 2) The female discriminant equation, when consideration was given to female students and the regression equation containing the additional social science variable, when consideration was given to male students, had a degree of effectiveness in identifying potential drop-out students. Further, it may be stated that the two equations cited were sufficiently sound statistically to be utilized by trained student personnel workers in the respective institutions. And 3) no single item of information available for freshmen students affords an adequate index to identify potential drop-out students, rather a constellation of selected traits and qualities are needed to assist in curtailing a student mortality rate from a prediction direction.

REFERENCES

1. Allen, Charles M. Combating the Dropout Problem (Chicago: Science Research Associates, Inc., 1956), 47 pp.
2. Carey, Matthew G. The Relations Between Selected Variables and the Cumulative Grade Average of Students Enrolled at the University of Wyoming, unpublished Ph. D. dissertation, University of Wyoming, 1955.
3. Iffert, Robert E. Retention and Withdrawal of College Students, U. S. Department of Health, Education and Welfare, Office of Education, Bulletin No. 1 (Washington, D. C.: United States Government Printing Office, 1958).
4. Iffert, Robert E. "Study of College Retention and Withdrawal," College and University, XXXI (Summer, 1956), pp. 435-77.
5. Wert, James Edwin, Neidt, Charles O., and Ahmann, J. Stanley. Statistical Methods in Educational and Psychological Research (New York: Appleton-Century-Crofts, Inc., 1954), 435 pp.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 7, March 1963)

Teacher Roles Before and After Internship

RUSSELL P. KROPP
JOHN E. ANDERSON, JR.
Florida State University

PROVISIONS for inter-teaching, student-teaching, or practice-teaching experiences are incorporated in teacher-education programs to give the prospective teacher an opportunity to "try-out," apply, practice, and perfect in a practical situation that which he has learned about teaching. Also, student-teaching is incorporated in programs, at least partly, for the opportunity it provides the intern to develop realistic attitudes about the role which a teacher must take in the classroom and in the public school context. The study presented here investigates the shaping of such attitudes about teaching during the period of internship.

Generally, the subsequent information deals with interns' conceptions of the ideal of the teacher (i. e., the most desirable role a teacher might take in a situation devoid of practical necessities and restraints), and the actual role of the teacher (i. e., the role which the teacher must take in a real teaching situation tinctured with whatever restraints, urgencies, and necessities the teacher might perceive). Changes in such perceived roles are studied over the internship period. Also, the relationships between such roles as perceived by the interns and the conceptions of such roles as perceived by the public school directing teacher and the college methods teacher, separately, are studied. In addition, the interns' ability to predict these conceptions of the directing teacher and methods teacher, separately, is studied.

Method

Subjects were all undergraduate seniors enrolled in the School of Education "Internship" block of courses. The internship block lasts for one semester and consists of eight weeks of full-time practice teaching under the direction of a public school teacher in the intern's teaching field, and a methods course, part of which precedes and part of which follows student teaching, which is taught on campus by a college instructor. There were 30 students enrolled in one such block. All were originally included in the study, but because of failure to turn the needed data in, either by the intern or his directing teacher, the number for whom complete records were obtained was only nine. All subjects were prospective secondary-school teachers. All had the same methods teacher, but none had the same directing teacher. Because all had the same methods teacher, and because complete sets of data could be obtained for only nine subjects, the results of the investigation must be regarded with extreme tentativeness.

The Instrument. Written statements descriptive of the various roles of the teacher were prepared and submitted to judges for evaluation as to relevance to teacher roles. Thirty statements were finally selected. Six statements were chosen for each of five categories representing the various roles the teacher is expected to play. The categories were:

1. student relationships
2. fellow-teacher interactions
3. supervisor and administrator relationships
4. community relationships
5. engagement in activities leading to his own professional development

Procedure

The data were collected by having the respondent, e. g., the student teachers or interns, the directing teachers, and the methods teachers, rate the 30 statements with regard to some criterion, i.e., the ideal role of the teacher, but with the condition that the rating had to be given a specific normalized distribution. The statements were to be assigned to seven categories: category "1" was to contain the most descriptive statement and category "7" was to contain the least descriptive statement. The number of statements to be assigned each category was as follows:

Statement Value	1	2	3	4	5	6	7
Number of Statements	1	3	6	10	6	3	1

The statements were sorted with regard to the interns' conception of the ideal role of the teacher. At the end of the semester only, they sorted the statements with regard to the actual role of the teacher; their prediction of the actual sorts of their methods teachers and their directing teachers, separately; and their prediction of the ideal sorts of their methods teacher and directing teachers. The directing teachers and the methods teachers were also asked to sort the statements at the end of the semester with regard to their conceptions of the actual and ideal roles of the teachers. In describing the data, "ideal" means "ideal role of the teacher," "actual" means "actual role of the teacher." From each of the groups, i.e., student teachers, directing teacher, and methods teacher, the following data were collected:

Student teachers:
- Student's ideal sort at beginning of semester = (SI1)
- Student's ideal sort at end of semester = (SI2)
- Student's actual sort at end of semester = (SA)
- Student's prediction of directing teacher's ideal sort at end of semester = (SPDI)
- Student's prediction of directing teacher's actual sort at end of semester = (SPDA)
- Student's prediction of method teacher's ideal sort at end of semester = (SPMI)
- Student's prediction of method teacher's actual sort at end of semester = (SPMA)

Directing teacher:
- Directing teacher's ideal sort at beginning of semester = (DI)
- Directing teacher's actual sort at beginning of semester = (DA)

Methods teacher:
- Method teacher's ideal sort at beginning of semester = (MI)
- Method teacher's actual sort at beginning of semester = (MA)

Method of Analysis. The correlations between sets of sorts were determined by the Q-sort technique. Since there were nine subjects, in each case, except when otherwise noted, nine correlations were obtained and the average correlation, according to the Fisher z-transformation, was computed and used as the descriptive statistic.

Results

This section is divided into four parts: 1) correlation between ideal sorts, 2) correlation between actual sorts, 3) correlation between ideal and actual sorts, and 4) correlation between predicted sorts. The information is presented in tabular form. The column titled "variables" indicated the two sorts for which the correlation was obtained. The column titled "$\bar{r}$" is the average correlation of the nine correlations between the variables. The column titled "Range" contains the lowest and highest correlation obtained between the pair of variables.

From the data in Table I, it is apparent that interns' conceptions of the ideal role of the teacher change very little during the semester in which they practice teach. Their conception of the ideal role at the end of the semester are slightly more in agreement than are their ideal roles in the beginning, with the ideal roles stated by the directing and methods teachers. Their similarity with their directing teachers is greater than that with their methods teacher, both in the beginning and in the end. The moderate correlation between the ideal sorts of the directing and methods teachers indicates a lack of uniformity on the part of the interns' teachers about the ideal role of the teacher.

Interns show little similarity with their directing teachers as regards the actual role of the teacher, and they show even less similarity with their methods teacher. The methods and directing teachers are slightly more in agreement than the interns are with either of them. Table III shows that the ideal and actual sorts of interns are highly similar. The agreement between end-of-semester ideal and actual is substantially greater than the agreement between beginning-of-semester ideal and actual. Interns are moderately successful in predicting the ideal sorts of their directing and methods teachers. The magnitude of their correlations are about the same size as those correlations in Table I between their ideal sorts. However, their predictions are much more highly related to their own ideal sorts than they are to those of their directing and methods teachers.

The interns are clearly less successful in predicting the way in which their directing teacher and methods teacher made their actual sorts. But again their predictions are more a function of their own conception of the actual role of the teacher than of

TABLE I

CORRELATIONS BETWEEN IDEAL SORTS

Variables	$\bar{r}$	Range
(SI1) (SI2)	.70	.54 to .87
(SI1) (MI)	.43	.18 to .65
(SI2) (MI)	.46	.22 to .72
(SI1) (DI)	.54	.26 to .80
(SI2) (DI)	.58	.33 to .68
(MI) (DI)	.47	.17 to .68

TABLE II

CORRELATIONS BETWEEN ACTUAL SORTS

Variables	$\bar{r}$	Range
(SA) (MA)	.16	.00 to .24
(SA) (DA)	.25	-.17 to .50
(MA) (DA)	.29	.04 to .52

TABLE III

CORRELATIONS BETWEEN IDEAL AND ACTUAL SORTS

Variables	$\bar{r}$	Range
(SI1) (SA)	.66	-.09 to .87
(SI2) (SA)	.85	-.04 to 1.00

TABLE IV

CORRELATIONS BETWEEN PREDICTED SORTS

Variables	$\bar{r}$	Range
(SPDI) (DI)	.49	.31 to .72
(SPDI) (SI2)	.69	.52 to .91
(SPMI) (MI)	.41	.02 to .54
(SPMI) (SI2)	.67	.20 to .85
(SPDA) (DA)	.28	-.02 to .55
(SPDA) (SA)	.52	.39 to .68
(SPMA) (MA)	.20	.07 to .42
(SPMA) (SA)	.56	.24 to .81

their knowledge of their supervisors' conceptions of the actual role.

It was found that 1) there was a high relationship between the beginning and end-of-semester ideal sorts of interns; 2) at the end of the semester their conception of their actual role was much like their conception of the ideal role; 3) their actual sorts were only moderately correlated with the actual sorts of their directing teachers and slightly less so with the actual sort of their methods teacher; 4) they predicted the ideal sort of their directing teacher better than they did the ideal sort of their methods teacher; 5) they predicted the directing teacher's actual role better than they did the methods teacher's actual sort; 6) all their predictions, whether of ideal or actual, were more in keeping with their own ideal and actual sorts, respectively, than they were with the sortings of the directing teachers or methods teachers. It also found that the directing teachers and the methods teachers were only on moderate agreement about the ideal role of the teacher, and that directing teachers and methods teachers were even in less agreement about the actual role of the teacher.

Discussion

The conclusions drawn above and the discussion here are indeed tentative in light of the few cases on which the study is based. First, the similarity between the actual and ideal sorts of the interns is somewhat heartening in that, as the interns see it, there is no substantial difference between what they ideally would like to do and what they actually can do. Second, the internship semester does not seem to be an especially fruitful time during which to develop attitude about the role of the teacher. Perhaps this is attributed, in part, to the intern's going into the situation with rather well-developed attitudes which are tenaciously held. Third, interns are not clearly aware of the actual role which the methods teacher or directing teachers advocate taking. This is revealed by the intern's lack of ability to predict the sorts of both the directing teachers and the methods teacher. It is difficult to believe that they were not informed of roles which were advocated inasmuch as the directing teacher supervised and corrected his work and the methods teacher advised him about what to do when teaching. Possibly interns unconsciously misinterpreted the roles that were advocated in order to make them fit their own conceptions. This is somewhat supported by the fact that their predictions were more highly related to their own conceptions of the role than to the conceptions they were attempting to predict. Fifth, interns might be in a problematic situation with regard to the actual role to take or the idealization of the role, inasmuch as the correlations indicate that there was only moderate agreement between the methods teacher and the directing teachers on the matter.

In addition, it might be generally said that this study indicates a great need for clearer communication among those responsible for teacher preparation and further investigation into how and at what stage of development teacher role attitudes are shaped.

Summary

A set of 30 statements dealing with the roles of the teacher was prepared. Nine senior-level teaching interns sorted their statements at the beginning and at the end of the semester in which they practice–taught in public high schools. The statements were sorted with regard to the interns' conception of the ideal role of the teacher. At the end of the semester only, they sorted the statements with regard to the actual role of the teacher; their prediction of the actual sorts of their methods teachers and their directing teachers, separately; and their prediction of the ideal sorts of their methods teacher and directing teachers. The directing teachers and the methods teacher were also asked to sort the statements at the end of the semester with regard to their conceptions of the actual and ideal roles of the teachers. It was found that:

1. There was a high relationship between the beginning and end-of-semester ideal sorts of interns.
2. At the end of the semester, their conception of their actual role was much like their conception of the ideal role.
3. Their actual sorts were only moderately correlated with the actual sorts of their directing teachers and slightly less so with the actual sort of their methods teacher.
4. They predicted the ideal sort of their directing better than they did the ideal sort of their methods teacher.
5. They predicted the directing teacher's actual role better than they did the methods teacher's actual sort.
6. All their predictions, whether of ideal or actual, were more in keeping with their own ideal and actual sorts, respectively, whan they were with the sortings of the directing teachers or methods teacher.

It was also found that the directing teachers and the methods teacher were only on moderate agreement about the ideal role of the teacher, and that directing teachers and methods teachers were even in less agreement about the actual role of the teacher.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 7, March 1963)

The Effect of Visual Imagery Upon Spelling Performance

LEON D. RADAKER
Butler, Pennsylvania

WORD RECOGNITION ability occupies a central position in the American schools as it is a primary vehicle of learning. Penty (4) has shown that 50 percent of all school drop-outs were "poor" in word recognition ability while 14.5 percent were considered as "good" in this area. It has been estimated that ten percent of the school population is deficient in word recognition.

A diligent search of the literature revealed that Rosseler (6), Pear (3), and Susukita (7) contended that the ability to create images could be increased by practice. The Wheelers (8), Harris (2) and Allport (1) have assumed that children who possessed the ability to create vivid images had a distinct advantage over those unable to perform this task satisfactorily. The author (5) has shown in previous studies on normal and retarded children that visual imagery could be improved through training. Such training led to significant improvement in the ability to reproduce unfamiliar word forms.

In this study it is hypothesized that children who receive training in the creation of images will score higher in spelling activities.

Design and Procedure

Subjects

The subjects were 60 pupils enrolled in the fourth grade in the Butler Area Elementary Schools in Butler, Pa. The children ranged in age from 8 1/2 to 10 1/2 years. Sex differences in numbers were negligible as there were 31 boys and 29 girls.

Design

To control the number of subjects appearing in each experimental group, 20 slips of paper were prepared for each of three categories. These were designated as Imagery Practice 2, Imagery Practice 6, and Control. For purposes of identification the groups were abbreviated to C, IP2, and IP6. As each subject was drawn in random order, a slip of paper was selected from an opaque bowl and assigned to the group indicated. In this manner three groups were constituted with 20 subjects in each.

Variables

The dependent variables in this study were two forms of the Stanford Achievement Test, forms M and N. These tests are designated by the publisher as equivalent forms. All scores were converted into grade equivalents. The tests were given April 15, 1959 and 1960; thus, an entire year elapsed between form M in third grade and N in fourth grade.

The independent variables were two levels of imagery practice. The lower level was given two sessions of practice while the upper level received six sessions. Each session continued 45 minutes. Formal practice was completed within a two week period. Training was completed by the Christmas vacation, 1959.

Two practice lists of 40 words each were prepared. These were designated as A and B. The words were selected on the basis of being unique in pattern, low in frequency, and relatively unfamiliar to the subjects as judged by a panel of fourth grade teachers. List A was presented to the lower level while the upper level received equivalent amounts of practice on A and B. The practice lists are presented in Table I.

The imagery practice groups were given training in creating images of the words on the practice lists. Each word was typewritten in pica type in the center of an unlined 3 x 5 inch card. Each subject had a pack of cards containing the words in sequential order. As each word was presented the subject scrutinized the word to note the letter sequence. The subject then closed his eyes and tried to arouse an image of the word in large, glossy, black letters on a white background. This simulated the words as they would normally appear in a textbook. It was suggested that the word be visualized as though it were projected on a large outdoor theater screen. Once the image was aroused the subject tried to stabilize the image and retain it as long as possible with a maximum objective of one minute for each.

Unstable images were stabilized in this manner. The subject was asked to imagine that the words were created from large metallic letters with holes punched in the top and bottom to accommodate fantasy nails. In this way it was possible to visualize unstable letters as being nailed securely to the theater screen. This suggestion appeared to facilitate word stability.

To improve retention it was suggested that the subject imagine fantasy paste being applied to the backs of the letters to glue the words in place; a fantasy floodlight was shown upon the words with such brilliance that the word could not fade until the subject wished the image to disintegrate. The subjects reported that this device was valuable in training retention.

Procedure

Training was administered to five in a group. To minimize growth and maturation factors formal training was completed within a period of two weeks. It was suggested to the imagery practice groups that they adopt this method for studying spelling lessons. As each word was given, the subject was to try to create an image of the word with photographic clarity. Each experimental group was segregated from the others during the practice sessions. The control was limited to six sessions of free play and social conversation to minimize motivational factors which may have been present. The subjects in each group were asked not to communicate with each other about the various experimental tasks.

After the spelling scores of the Stanford Achievement Tests had been converted to grade equivalents, the data were subjected to statistical treatment by an analysis of covariance.

Results

Reliability, means, ranges, and standard deviation of the dependent variable

An estimate of the reliability of the dependent variable was provided by the initial and final test correlation of the control group. These are true reliability coefficients as they received no experimental treatment. The coefficients obtained for each group are shown in Table II. The minimum significant value for 18 df at the .05 level is .444 and at the .01 level is .561. Since all coefficients exceed these values, the reliability can be accepted as adequate.

The remaining coefficients for IP2 and IP6 are not true reliability coefficients because the extent to which the treatment effects scrambled the rank orders could not be predicted. The correlations, however, are quite high in view of these effects.

Table III contains the ranges, means, and standard deviations for the initial and final grade equivalents obtained on the standardized tests.

These findings suggest that imagery practice resulted in fairly uniform improvements since the mean values increased without greatly disturbing the rank orders of the subjects.

Comparisons testing the hypothesis

The data was subjected to three tests to determine the appropriateness of the covariance design. The tests used were the (1) significance of regression, (2) Bartlett's test of homogeneity of the variance, and (3) homogeneity of regressions. All tests supported the use of the analysis of covariance.

The hypothesis stated that children who received training in the creation of visual images would make significant improvements in spelling activities. This was tested by the grade equivalent spelling scores on the Stanford Achievement Tests, forms M and N, given at an interval of one year. Three comparisons were made to test this hypothesis: (1) C and IP2, (2) C and IP6, and (3) IP2 and IP6. The F-ratios are shown in Table IV. The minimum significant value for 1, 57 df at the .05 level is 4.01. The hypothesis was confirmed for both levels of imagery training. No significant differences appeared between the two levels of practice.

Discussion

It has been shown by the author in previous studies that imagery is effective in improving memory for abstract designs, objects, and word forms. These tasks were measured over a two week period.

In this study it was contended that the effects were beneficial over a longer period of time. In this study a period of one year elapsed between the initial and final tests. It is noted that the lower level of training was practically as effective as the higher level. This is consistent with previously obtained results. Apparently imagery can be trained in a relatively short period. This economy in time is an acceptable feature from an instructional point of view.

Imagery training probably permitted the subject to obtain sharply defined images of words which served as models in assessing similar representations. These models enabled subjects to discriminate words effectively. It is assumed that vividness of the image tended to fixate the words giving the subject greater confidence in recall.

It is suggested that vividness of imagery forcibly divorces figure from ground. Subjects reported that practice produced images that were clearer and sharper in detail than the original images.

Conclusion

It concluded that imagery is successful in improving spelling performance over longer periods of time. In this study a period of one year elapsed between the two testings. The improvement in spelling suggests that it will also be effective in reading because of the effort devoted to word recognition and recall.

REFERENCES

1. Allport, Gordon. "Eidetic Imagery", British Journal of Psychology, 15: 99-120, 1924.
2. Harris, Albert J. How to Increase Reading Ability, Longmans Green and Co. New York, 1947.
3. Pear, T. H. "The Place of Imagery in Mental Processes", Psychological Abstracts, 12:729

1938.
4. Penty, Ruth C. "Reading Ability and High School Drop-outs", Contributions to Education, Teachers College, Columbia University, 1956.
5. Radaker, Leon D. Visual Imagery as a Factor in Memory for Word Forms, Unpublished doctoral dissertation, Pennsylvania State University, 1959.
6. Rosseler, F. "Distribution and Appearance of Subjective Optical Eidetic Images in Boys and Girls from Six to Ten Years Old", Psychological Abstracts, 3: 4741, 1929.
7. Susukita, T. "On Memory Images of a Perceptual Character", Psychological Abstracts, 11: 5022, 1937.
8. Wheeler, Lester R. and Wheeler, Viola D. "Dyslexaphoria: Symptoms and Remedial Suggestions", Elementary English, 32: 305-311, 1955.

TABLE I

IMAGERY PRACTICE WORDS

	List A		
zebu	upas	topaz	uncial
dais	tolu	nymph	phylum
ywis	toga	xylem	syrinx
weft	thegn	tuque	syntax
zany	voile	uvula	swaraj
wadi	umbra	viand	striae
visa	uhlan	virtu	sirdar
lieu	tibia	whelk	tautog
orgy	triad	gypsy	thymus
widu	myrrh	xebec	guelph
	List B		
bias	prism	bantu	bryony
cyme	adieu	culex	wapiti
acne	mauve	kazak	phloem
lyre	tiara	feaze	concha
lynx	axial	kauri	drachma
pygmy	bayou	benign	corymb
sylph	align	biceps	deodar
abyss	aorta	bisque	cuisse
crypt	algid	milieu	scherzo
feign	helix	boucle	pleiad

TABLE II

COEFFICIENTS OF CORRELATION FOR FORMS M AND N

C	.77
IP2	.70
IP6	.78

TABLE III

RANGES, MEANS, AND STANDARD DEVIATIONS FOR STANFORD ACHIEVEMENT TESTS M AND N

	Initial			Final		
Group	Range	Mean	SD	Range	Mean	SD
C	3.2-6.6	4.43	.86	3.5-6.9	5.02	.82
IP2	2.2-6.0	4.15	.66	4.3-6.9	5.17	.61
IP6	3.4-5.3	4.33	.25	4.3-6.9	5.37	.40

TABLE IV

F-RATIOS TESTING THE HYPOTHESIS

Comparison	F-Ratio
C and IP2	5.05*
C and IP6	5.18*
IP2 and IP6	.06

* significant at .05 level

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 7, March 1963)

An Evaluation of the Vocational Counseling Program of the Employment Security Service in Kanawha County

WILLIAM V. WAGNER
West Virginia University

THE PROBLEM of placing workers in suitable jobs has become more imminent with the continued industrialization of the United States. This has become an increasing concern of high school graduates with little or no work experience. The United States Employment Security Service has made an effort, through its employment counseling program, to meet this need.

The object of this study was to evaluate the effectiveness of employment counseling services in Kanawha County, West Virginia. The major purpose of the study was to ascertain strengths and weaknesses of the program and make recommendations for improvement.

Procedures

The investigation included three groups of subjects: 1) a sample of 1957 high school graduates from 11 Kanawha County, West Virginia high schools who received employment counseling services, 2) a sample of 1958 high school graduates from eight Kanawha County, West Virginia, high schools who were counseled and tested by the employment counselor, and 3) senior teachers, counselors, and principals of the eleven high schools participating in employment counseling services.

A questionnaire was devised and mailed to 391 graduates included in the 1957 follow-up, approximately eighteen months following high school graduation. A total of 202 graduates responded, and data were tabulated and analyzed using the chi-square technique, ranking method, or percent of appropriate group. The 1958 follow-up included 124 graduates who completed a questionnaire during a group interview session immediately following graduation. Data were summarized into related areas and interpreted using the chi-square technique.

A check list was developed and mailed to school principals, counselors or deans, and senior teachers to sample the reaction of school personnel regarding employment counseling services. Fifty-nine responses were received, tabulated, and analyzed using the ranking method and percentage distribution.

Summary of the 1957 Follow-Up

Results of the investigation were summarized into two major groups, those working full-time at the time of the follow-up and those who had worked but were not employed full-time at the time of follow-up. There were 32.18 percent of the working group and 23.37 percent of the non-working group who had gained employment in line with their vocational plan. The largest percentages who had worked in line with their plan were employed in clerical or sales work. The primary reason given by both groups for not following their vocational plan was lack of job opportunities. A significant number of graduates from both groups had held two jobs since graduation and would not select the same occupational goal if formulating a new vocational plan. They credited parents, teachers, friends, and relatives as most influential in helping them decide on an occupational goal before graduation, ranked tests as most helpful in arriving at a vocational plan, gave reasons other than alignment with the vocational plan for accepting jobs, had not sought help from the employment counselor since graduation, had not been contacted by the employment office since graduation for job referrals, would recommend employment counseling to friends, and indicated the eleventh grade as the most helpful time for employment counseling. More members of the working group had made progress in their jobs, expressed satisfaction with their last job, desired to remain in their last jobs, were satisfied with pay received, had received more pay raises, were more secure in their last employment, and had discussed job interview techniques with the employment counselor.

Summary of the 1958 Follow-Up

Graduates stated that the counseling session was well planned and structured, counseling strengthened their self-understanding, tests were properly administered and interpreted, and the group orientation session was valuable in creating an interest in vocational planning. They were undecided about their life's work before counseling. Sufficient information was gained during counseling for vocational planning purposes. Graduates were not influenced by the employment counselor in their educational planning during high school. They agreed they would recommend the service to fellow students, felt additional interviews with the employment counselor would be helpful, planned to seek help through the employment office in finding a job, and stated employment counseling would be of most value during the tenth year in school. Occupational information discussed and considered with the employment counselor yielding significant responses, included the future of the occupation, job demand, location, opportunities, and general educational qualifications. Information discussed and yielding inconclusive ratios of significant results due to negative responses were the lack of consideration of working conditions, dangers or hazards, special tools or licenses needed for job entrance, practices of discrimination, retirement policies, fringe benefits, a description of the work, number of workers engaged in the occupation, the size of the employing business, stability of the job, average beginning wage, normal working hours, and requirements for entrance.

Summary of the School Personnel Check List

At the time of reply, school personnel were most familiar with the testing services offered in their schools by the employment counselor. Bulletins were received or meetings were attended explaining the purpose and extent of the services to approximately one-half of the school personnel. A significant number of school personnel reported that employment counseling services were of great value to students in meeting student needs and stimulating interest in more serious future planning. Of those school personnel responding, three in five expressed a willingness to assist the employment counselor if needed. A significant number of school personnel were doubtful or stated that homeroom teachers with the assistance of school counselors or deans were unable to provide services offered by the employment counselor and obtain the same results. If given personnel, materials, and time; three in five indicated they could provide more adequately the services offered by the employment counselor. School personnel agreed that school time should be provided for employment counseling, but were doubtful or stated sufficient time was not being provided. Three in five of the school personnel reporting stated that employment counseling results were made available to them for working with students. School personnel would recommend the service to non-participating schools and ranked the twelfth grade as the most helpful year for employment counseling.

Conclusions of the 1957 Follow-Up

1. The largest percentage of workers from both groups had formulated vocational goals related to clerical or sales work and had been employed in that classification. The major reasons given by both groups for not following their vocational plan was a lack of job opportunities. The percentage of graduates who had worked in line with their vocational plan was low in comparison with those studies reported in the literature.

2. There was no difference in the number of jobs held since graduation between the working and non-working groups. Both groups secured jobs through their own efforts or with the help of friends and relatives, and overestimated the positive relationship existing between job and vocational plan. A significant number of both groups would not select the same occupational goal if they were working out a new vocational plan.

3. Parents, teachers, friends, and relatives were the major influences in helping students decide on an occupational goal before graduation. Both groups stated that aptitude testing received major emphasis and was the most helpful service offered by the employment counselor for vocational planning purposes.

4. Graduates placed minor importance on the vocational plan when seeking employment.

5. Approximately three in five of the graduates of both groups were uncertain or stated they would be easy to replace in their last jobs, but indicated they were capable of more advanced positions.

6. A significant number of both groups had not requested help in finding work from the employment office since graduation. More of the working group than the non-working group had been contacted by the employment office since graduation for job referrals and counseling, an important factor relating to proper job placement.

7. An understanding of interview techniques may be a factor in accounting for a high percentage of the working group gaining employment in line with their vocational plan.

8. While working and non-working groups were equally enthusiastic about recommending employment counseling services to friends, there is a marked discrepancy between graduates' opinions and their actions in recommending employment counseling services to friends.

9. The working group gained more from counseling than the non-working group and saw an opportunity to benefit more from counseling received earlier in the high school program.

Conclusions of the 1958 Follow-Up

1. Although proper scheduling of a definite time for the interview was not amply provided, the counselor made adequate preparation for the interview and created an emotional atmosphere conducive to proper counseling procedures.

2. Graduates gained insight into vocational planning and additional self-understanding through employment counseling.

3. Responses of graduates were highly significant, indicating that test administration and interpretation were effective. Before counseling, an equal number of graduates had either underestimated or overestimated their abilities to do a particular line of work.

4. Graduates were undecided on an occupational goal prior to counseling.

5. Graduates indicated that occupational information was made available and discussed with the employment counselor, although many areas pertaining to job information were not considered. This was a factor distracting from the vocational plan as an aid in seeking employment.

6. The group orientation session motivated and emphasized to students the value of employment counseling, as an aid to vocational planning.

7. Graduates were not influenced by employment counseling in planning their educational programs. This was due to counseling services received after the last semester of the senior year was in progress. After employment counseling, they felt a need for additional education to meet vocational objectives.

8. A significant number of graduates stated they would recommend the service to fellow students.

9. Graduates indicated that additional interviews with the employment counselor would be of value, and they would seek help in finding employment from the employment office.

10. Immediately following graduation, graduates stated that the employment counselor's services would be of most value during the tenth year and of least value during the twelfth year.

Conclusions of the School Personnel Check List

1. School personnel most familiar with employment counseling services ranked in order were: 1) school principals, 2) school counselors, 3)senior teachers, and 4) senior homeroom teachers. Employment counseling services most familiar to school personnel ranked in order were: 1) tests, 2) vocational counseling, 3) occupational information, and 4) placement and follow-up.

2. Principals and counselors were better oriented regarding employment counseling services offered in their school than were senior teachers.

3. School personnel reported employment counseling services were of great value in stimulating students to give more serious consideration to future needs.

4. About three in five of the school personnel were willing to assist in the employment counseling program if needed. Slightly over one-half of school personnel felt there was a coordinated effort existing between the school and employment counselor in providing counseling services.

5. School personnel stated they could not provide effective employment counseling services to students without the help of the employment service, unless given adequate materials, time, and qualified personnel.

6. Respondents specified that school time should be made available for employment counseling and that ample time to consider all available data was not provided.

7. School personnel disclosed that all services offered by the employment counselor should be provided.

8. Results of employment counseling were more readily available to principals and counselors than senior homeroom teachers and senior teachers.

9. School personnel would recommend the employment counseling services for high school youth.

10. School personnel specified that employment counseling services were of most value during the twelfth grade.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 7, March 1963)

A Longitudinal Study of Auditory Discrimination *

BERTHA BOYA THOMPSON
Talawanda High School
Oxford, Ohio

THE PURPOSE of this study was to determine the relation of auditory discrimination (AD) and intelligence test scores to success in primary reading. Secondary problems were: 1) to determine whether children made significant improvement in auditory discriminative ability while attending the first and second grades; 2) to determine whether poor readers in primary reading establish a pattern on the twelve subtests of the Wechsler Intelligence Scale for Children (WISC) which might be significantly different from a pattern established by good readers.

Method and Procedure

Most past studies which dealt with the relationship of auditory discrimination and success in reading were either cross-sectional or measured success in reading at the close of the first year. This present study differed in that it was longitudinal. All the first-grade children who entered the two elementary schools of Oxford, Ohio, in September, 1958, were tested. The last test was administered in May 1960, to 105 children when the group was completing second grade.

Tests that measured auditory discrimination, intelligence, reading achievement, and auditory acuity were administered. The three tests of auditory discrimination were: A Test for Auditory Discrimination, Form A, by Wepman; Boston University Speech Sound Discrimination Picture Test; and Auditory Discrimination and Orientation, a subtest of the SRA Reading Analysis: Aptitude, Form A. The directions of the latter subtest were modified to meet the needs of this young age group. These three tests were administered to the sample in the month preceding entrance to the first grade and were readministered in the eighth month of the second grade. A composite AD score was used in this study.

The Wechsler Intelligence Scale for Children was selected because it provided three measures of intelligence. Mental ages could also be derived from the Full Scale IQ's. Moreover, the mean scores on the WISC subtests may be analyzed for patterns of such groups as poor readers and good readers further establishing the relation of intelligence to reading. The twelve WISC subtests were administered when the mean chronological age of the sample was seven years, four months or during the month prior to the entrance of this group to the second grade. Two measurements of reading achievement were used: Gates Advanced Primary Reading Test, Type AWR, Form 1, and Type APR, Form 1. The former measured word recognition, the latter paragraph reading. These tests were administered when the sample was in the ninth month of the second grade.

Individual pure-tone audiometer tests were given to those who failed the preliminary screening tests. These tests were administered when the sample was in the eighth and ninth months of the second grade.

An arbitrary scaling method was used to determine the adequacy of the auditory discriminative ability of each individual. This scaling method was based upon research made by Wepman on his newer (1958) AD test.

Good readers and poor readers were established using two criteria. Criterion one included those 24 who placed at the upper end and those 24 who placed at the lower end of the reading distribution, using the composite reading scores. Criterion two included those 24 whose reading ages (R. A.'s) most greatly exceeded their mental ages (M. A.'s), and those 24 whose mental ages were higher than their reading ages.

Conclusions and Recommendations

The following conclusions are based upon findings from the analysis of data of this study using the seven different tests and this small total school population. Some are concluded by implication from these findings:

1. From the statistical analysis of the data of this study, it would appear that auditory discrimination and intelligence are highly correlated with

---

* For a complete description of this research, see The Relation of Auditory Discrimination and Intelligence Test Scores to Success in Primary Reading by Bertha Boya Thompson, Doctor's Thesis, School of Education, University, Bloomington, 1961), 138 pp.

success in primary reading.

2. From the high intercorrelations of the factors of auditory discrimination and intelligence, it may be concluded that adequacy in the one trait is frequently accompanied by adequacy in the other at the beginning of the first year of school.

3. From the performance of this sample on the three tests of auditory discrimination, it would appear that children grow or improve in auditory discriminative ability from the time that they enter first grade.

4. Inaccurate discriminative ability is more characteristic of first-grade entrants than accurate ability. The reverse is true at the end of the second grade.

5. Auditory discriminative ability is not fully developed in all second-grade students. Approximately 24 percent of this sample had inaccurate auditory discriminative ability at the end of the second grade. Approximately one-half of these were classified as poor readers. This again points towards the importance of adequacy in auditory discrimination to success in primary reading. However, it definitely reveals that other factors are important in reading disability.

6. The status of the first-grade entrant's auditory discriminative ability is highly prognostic in determining who will become a good reader. Of those 24 children who became good readers according to criterion one, 16 had adequate ability at the entrance to first grade. Of those who became poor readers, only one had achieved adequacy. However, the predictive value lessens when trying to determine whose R. A. will be higher than his M. A., for only twice as many who became good readers according to criterion two had accurate discrimination at the entrance to first grade as those who became poor readers.

7. The poor readers made greater proportionate gains than the good readers in auditory discrimination. The poor readers on both criteria exceeded frequencies expected at the close of the second grade; the good readers did not achieve expected frequencies. Nevertheless, there were more good readers than poor readers according to criterion one, with adequate ability at both testings. The opposite was found in the final testing data when criterion two was applied, but this difference was not significant. Gains were made by all groups, thus indicating the general development of this ability in most primary children.

8. There was a significant difference between the performance of good and poor readers on the initial and final distributions of auditory discrimination test scores when criterion one was applied, but not when criterion two was applied; therefore, high auditory discrimination test scores are more likely characteristic of good readers at the upper end of the reading-achievement distribution but not as likely when these good readers are those whose R. A.'s exceed their M. A.'s.

9. From the performance of the good readers and poor readers of criterion one on the WISC subtests, it can be concluded that intelligence is highly prognostic in determining who will become a good reader, placing at the upper end of the reading achievement distribution for the means of these good readers were significantly higher on ten of the subtests at the one percent level and on the Digit Span subtest at the five percent level. Only on the Coding subtest was there no significant difference.

10. From observing the behavior of this sample on the Coding subtest, it seemed to be greatly affected by motor coordination or facility in using a pencil.

11. From the performance of the good readers and poor readers of criterion two on the WISC subtests, it can be concluded that intelligence is not highly prognostic in determining who will or will not achieve an R. A. above his M. A., for although the poor readers scored higher means on every subtest, the difference was significant at the one percent level on only two tests, Similarities (Verbal Scale) and Picture Arrangement (Performance Scale). The mean differences on two other subtests of the Performance Scale: Picture Completion and Object Assembly were significant at the five percent level, however.

12. The mean Performance Scale IQ of the poor readers of criterion two was significantly higher than their mean Verbal Scale IQ. The difference between these two means of the good readers was not significant, nor did they differ significantly from the Verbal IQ mean of the poor group. Reading is a highly verbal skill; the lower correlations found in this study were those in which Performance Scale IQ's were used. Perhaps the reason these poor readers failed to reach an R. A. higher than their M. A. was this significantly high mean Performance Scale IQ.

13. The rank order of the mean WISC subtest scores varied greatly for each good reader group and its corresponding poor reader group.

14. The rank order of the mean WISC subtest scores of the good readers and poor readers of both criteria varied greatly.

The following recommendations were based upon the reviewed conclusions and their implications. Some of these are suggested changes in educational practice; others pertain to suggested research relative to the primary grades.

1. That first grade entrants be tested, using a reliable test of auditory discrimination. That those children who score at the lower end of such a distribution be given an extended readiness period in which exercises to develop this ability are presented.

2. That those who develop slowly in auditory discriminative ability be introduced slowly to phonics as well.

3. Those first-grade entrants who test at the upper end of a distribution of auditory discrimina-

tion test scores should be introduced to an accelerated program of word study; not, however, before an understanding of "reading for meaning" is developed.

4. That a similar study be made investigating the relation of auditory discrimination and spelling. That the status of auditory discrimination be measured at the time when spelling instruction is begun. That success in primary spelling be determined by achievement tests at the end of the third grade.

5. That a similar study be made investigating the relation of visual discrimination, measured at the beginning of first grade, and success in primary reading as measured by reading achievement scores at the end of the second grade. That the criteria for adequacy be established to determine whether visual discrimination is an ability in which a high level of performance is reached by most primary children.

6. That the same type of study that was made of subtest mean differences on the WISC of good and poor readers be made for those who had adequate or inadequate auditory discriminative ability at the beginning of the first grade.

7. Since there is evidence that the difficulty of the WISC subtests is most suitable for the middle age ranges - ten and a half to thirteen and a half, there may be some value in retesting this sample at age ten and a half and correlating the test results with the reading achievement and auditory discrimination scores of this study.

8. That the diagnostic significance of an objective measure of intratest scatter on the WISC subtests be investigated for primary children.

REFERENCES

1. Altus, Grace T. "A WISC Profile for Retarded Readers," Journal of Consulting Psychology XX (April 1956), pp. 155-56.
2. Bond, Guy L. The Auditory and Speech Characteristics of Poor Readers, Teachers College, Columbia University (1935), 48 pp.
3. Croley, Hugh T. A Statistical Study of the Wechsler-Bellevue Pattern for Poor Readers, Master's thesis, The Department of Psychology, The University of Denver, Denver (1949) 36 pp.
4. Durrell, Donald D., and Harrington, Mary James, Sister. "Mental Maturity versus Perception Abilities in Primary Reading," The Journal of Educational Psychology XLVI (October 1955), pp. 375-80.
5. Durrell, Donald D. and Murphy, Helen A., "The Auditory Discrimination Factor in Reading Readiness and Reading Disability," Education LXXIII (May 1953), pp. 556-60.
6. Graham, E. Ellis. "Wechsler-Bellevue and WISC Scattergrams of Unsuccessful Readers" Journal of Consulting Psychology XVI (August 1952), pp. 268-71.
7. Mary Nila, Sister. "Foundations of a Successful Reading Program," Education, LXXIII (May 1953), pp. 543-55.
8. McLean, James. "A Review of the Kansas Studies of Sound Discrimination," mimeographed copy of speech delivered to American Speech and Hearing Association at 1959 convention, 9 pp.
9. Reidy, Mary Elizabeth. A Validity Study of the Wechsler Intelligence Scale for Children and its Relationship to Reading and Arithmetic, Master's thesis, Department of Psychology and Psychiatry, Catholic University of America, Washington, D.C. (1951), 26 pp.
10. Smith, Donald E. P. and Carrigan, Patricia M., The Nature of Reading Disability (New York: Harcourt, Brace, and Company, 1959) 149 pp.
11. Templin, Mildred C. "A Study of Sound Discrimination Ability of Elementary School Children," The Journal of Speech Disorders VIII (March 1943), pp. 127-32.
12. Templin, Mildred C. Certain Language Skills in Children: Their Development and Interrelationships, (Minneapolis: The University of Minnesota Press, 1957), 183 pp.
13. Wepman, Joseph M. "A Test for Auditory Discrimination," 1941 (mimeographed).
14. Wepman, Joseph M. "Relationships of Auditory Discrimination to Speech and Reading Difficulties," A Journal of the American Speech and Hearing Association (abstract of speech), I (November 1959) 96 pp.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 7, March 1963)

Degree of Teacher Domination in Group Processes and Student Attitude Toward Instruction in Secondary Methods Classes*

KENNETH H. HOOVER
Arizona State University

CONTINUED EFFORTS have been made to determine the relative effectiveness of different teaching methods upon achievement of college students. Among these are studies by Birney and McKeachie (1), Guetzkow (2), Lyle (3), and Stovall (7). For the most part, such studies reveal that the method of presentation tends to have little or no bearing upon academic achievement. When other types of outcomes are considered, however, student-centered approaches often tend to be favored slightly. Through student-centered instructional programs, such as those of Birney and McKeachie (1), Patton (5), Wispe (8), and Wispe (9), students apparently develop more tolerant attitudes, gain a greater sense of initiative and independence, improve in certain emotional and social skills, and enjoy the student-centered sessions more. The present study was an attempt to determine the relative influence of teacher-centered, pupil-centered, and group-centered instruction upon attitudes toward both the instructor and the course involved.

Procedure

The study involved college seniors in a four-semester credit course, "Methods of Secondary School Teaching," taught by the investigator. Classes met for two-hour sessions twice each week for an entire semester. Each of three classes (approximately 30 per class) was taught by purely teacher-centered methods for approximately 12 weeks during the 1960-61 school year. Attitudes of the groups toward certain aspects of the instructional program were then determined by the administration of the Purdue Rating Scale for Instruction (6). For the remaining six weeks of the term, one of the groups was subjected to a teacher-centered approach, a second group was directed by a pupil-centered approach, while a third class was purely group-centered in instructional approach. At the conclusion of the course, the Purdue Rating Scale was re-administered to the three groups and comparisons of the attitude changes were made by means of analysis of variance. For the purpose of clarification, the different instructional approaches have been described in the section which follows.

Instructional Approaches Employed

Teacher-Centered Approach: This approach was characterized by a variety of teaching methods and techniques. Students gained experience in hearing (and in many cases participating in) lectures, oral and written reports, films, class and panel discussions, debates, and sociodramas. In all cases, however, the instructor determined the method to be used and when it would be applied. Assignments were made and tests were given as the instructor deemed appropriate.

Pupil-Centered Approach: The method employed in this approach is known generally as teacher-pupil planning (4). Students were asked to identify their own problems by submitting questions in the desig-

---

* A paper presented at the Annual Meeting of the American Educational Research Association, Chicago, Illinois, February 22-25, 1961.

TABLE I

SUMMARY OF THE ANALYSIS OF VARIANCE

Source of Variation	df	Sum of Squares	Means of the Squares	f ratio
Degree of teacher-domination	2	132	66	1.86
Aspect of the course	1	27	27	<1
Interaction	2	102	51	1.45
Within	159	5603	35.2389	—
Residual (interaction + within)	161	5705	35.4348	—
Total	169	5864		

nated area of study. With the help of the instructor, problems were grouped into areas which were then explored by interested groups of students. Each group was responsible for reporting its findings to the class. Manner and style of reporting were left to those involved, except that reports were to be judged by standards previously established by the entire class. Overall evaluational procedures of the unit also were determined by the entire class group. Techniques of reporting actually applied included panel discussion, debate, films, and committee reports. The instructor's role was to suggest, recommend and generally guide learning activities.

Group-Centered Approach: In this approach, emphasis was placed upon the importance of students working together as a dynamic group. Chairs were arranged in a large circle. A class leader, recorder and three group observers were elected and retained until the group decided to replace them with other students. To expedite the instructional process, a steering committee was elected for the purpose of recommending areas of investigation and procedural matters of all kinds. At the end of each class period, class observers discussed group problems briefly. They also made individual evaluations on the basis of criteria established by the group. Instructional techniques were determined by the group. In addition to round table discussion, the class made use of individual reports. The instructor served as the fourth class observer and as a member of the steering committee. Although he remained in the classroom during the entire time, he did not join the group during the discussion periods. After the termination of each discussion period he added, clarified, or sometimes questioned points of the discussion. This covered a period of 5 – 15 minutes of the 100 minutes of actual class time.

The results of the instructional approaches employed are given in Table I.

Summary and Conclusions

The present study involved the use of group processes in three sections of a college course in "Methods of Secondary School Teaching." All students held senior or graduate standing. An attempt was made to determine the effect of different degrees of teacher-centeredness upon student attitudes toward the instructor and the course. Attitudes were measured by making use of the Purdue Attitude Scale for Instruction before and after the experiment. A two-way analysis of variance indicated no significant differences in any of the variables studied.

REFERENCES

1. Birney, Robert, and McKeachie, Wilbert J. "The Teaching of Psychology: A Survey of Research since 1942," *Psychological Bulletin*, LII (January 1955), pp. 51-68.
2. Guetzkow, Harold S., et al. "An Experimental Comparison of Recitation, Discussion, and Tutorial Methods in College Teaching," *Journal of Educational Psychology*, XLV (1954), pp. 193-207.
3. Lyle, Edwin R. "An Exploration in the Teaching of Critical Thinking in General Psychology," *Journal of Educational Research*, LII (December 1958), pp. 129-133.
4. Parrish, Louise, and Waskins, Yvonne. *Teacher-Pupil Planning* (New York: Harper and Brothers, 1958), 159 pp.
5. Patton, Joseph A. *A Study of the Effects of Student Acceptance of Responsibility and Motivation on Course Behavior*, doctor's thesis (University of Michigan: Ann Arbor 1955), 168 pp. Abstract: *Dissertation Abstracts* 15: 637-638, N. 4 (1955).
6. Remmers, H. H., and Elliott. *The Purdue Rating Scale for Instruction* (Lafayette, Indiana, 1950).
7. Stovall, Thomas F. "Classroom Methods: Lecture vs. Discussion," *Phi Delta Kappan*, XXXIX (March 1958), pp. 255-258.
8. Wispe, Lauren G. "Evaluating Section Teaching Methods in the Introductory Course," *Journal of Educational Research*, XLV (1951), pp. 161-186.
9. Wispe, Lauren G. "Teaching Methods Research," *American Psychologist*, VIII (1953), pp. 147-150.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 7, March 1963)

Evaluating Teaching Personality Before Student Teaching Begins

LEBERN N. MILLER
Trenton State College
Trenton, New Jersey

HOW MANY supervising teachers, upon encountering a student teacher whose personal qualities or attitudes toward teaching and dealing with children do not seem to equip him for his teaching assignment, have remarked or secretly observed that the condition should have been brought to light earlier and the student counseled out of teacher preparation. It is tragic and wasteful for such a student to spend his and others' time until the last year of college, then to be told that he will not be recommended for a teaching position, or that he will not be permitted to finish his college work. It is worse and more tragic for college officials to allow him to slip by, to enter the profession, to add to the number of unhappy and unfit teachers.

Recently an elementary principal, upon being asked how colleges could be of greater service in providing good classroom teachers, replied that more should be done to improve the "teaching personalities" of students who plan to teach. If such an improvement program is not possible, he added, the least that college officials should do is to begin early, around the sophomore year if not earlier, evaluating students' teaching personalities and screening out or redirecting on the basis of sound evaluations.

Researchers have tried to arrive at a dependable procedure for predicting future success in teaching, sometimes by using students' scores on standardized or objective-type tests. The results have not been completely satisfactory, to say the least. Many such personal inventories showed little better than random predictive capability. Important elements, which appear to be untouched by objective devices, are the individual's teaching personality traits and attitudes. Undoubtedly these elements play a large part in the teacher's success in dealing with and developing the potentials, attitudes, and personalities of youngsters. These qualities and their evaluation, under present conditions, call for subjective inquiry and judgment, under conditions which are as completely controlled as possible.

The present study was an attempt to find out what is being done now, in the teacher preparation institutions of the nation, to identify and redirect students whose personality traits seem to limit their chances of success in the profession, as has been described. The writer assumed that in order for such evaluation to take place, the college must be carrying on an "experience-type" course for freshmen and/or sophomores. Using a questionnaire, the writer contacted 96 colleges throughout the nation and found that at 44 of them, such a performance-type course was being given at either or both of these levels. He then contacted the 44 schools and inquired as to what was being done in the way of evaluating the students' personalities, identifying and redirecting "poor risks" in the profession. Replies were received from 30 of the colleges. This is a summary of the data obtained.

Evaluative Programs in Thirty Colleges

Twenty-five of the 30 colleges reported that such a personality evaluation program was being carried on. One of the other five respondents questioned the wisdom of such a program in view of the lack of validity of present evaluative instruments, which will be seen later to be annoying to several others in the survey. At 16 of the 25 colleges, the evaluation was described as a matter of formal and written policy. Nine stated that the policy was not in written form, and seven declined to describe the procedure as formal. Two respondents added that their procedures were both informal and formal, that instructors sometimes report questionable cases without going through channels.

At nine of the 25 colleges having such programs, written evaluations of personal traits and attitudes were being filled out for each student in professional courses for freshmen and sophomores. Ten others indicated that written evaluations were made out only for individuals showing questionable behavior. Probably one may assume safely that at many colleges, evaluations are given orally; that was reported as being the case at one of the colleges.

There was considerable uniformity as regards the person doing the evaluating and identifying of possible cases. That was the instructor of the professional experience course in 21 colleges. In a few instances, such evaluation was being done by more than one person, by the social agency director, in three colleges where such agencies cooperated with the college in the experience, and by the dormitory supervisor or adviser (two colleges).

The evaluations having been done, continued the inquiry, to whom are they submitted? Most frequently, in seven reporting colleges, the recipient is the dean or chairman of the education department. At three other colleges, the evaluations were being submitted to the Committee on Admissions and Standing, also known as the Committee on Selection and Retention. At two colleges, the recipient was the coordinator of the course, at two others the Pre-Teaching Counselor, and at two others the Student Personnel Department.

The question was asked, "Who makes the final decision as to the disposition of a student case under question?" This important decision most commonly came from a committee or group, and judging from the comments, followed a great deal of thought and debate. The final arbiter in eight colleges was the Committee on Admissions and Standing. Other final arbiters, and the frequency in colleges so reporting, were as follows: the college president (three colleges); the faculty of the department of education (two colleges) and of education and psychology (two colleges); the departmental faculty plus the dean (two colleges; the individual counselor (two); the coordinator of student personnel (one college); and a committee composed of the chairman of education, the dean, the student's advisor, and a faculty member from the student's major field (one college). At two colleges where provision was made for appeal, the Committee on Appeals of course had the final responsibility.

To determine the real strength or effect of the programs, the writer asked whether a student actually could be dropped from school or redirected to another field on the basis of the evaluations. Sixteen answered affirmatively; four said no. One person said that further education courses were denied the student; another said the student was strongly advised as to his vocational future and that the results had been very effective. Two stated that the student could be redirected but not dropped; another that he is redirected if at all possible. One ventured the judgment that the commitee was too cautious in dealing with these cases.

The investigator tried to find out the numbers of such "poor risks," over a five-year period, that had been identified, redirected, and screened out. Vague responses were given, perhaps because of inadequate records or reluctance to disclose figures, but the information gained is presented here:

a) The reported number discovered over the past five years and the frequencies were: less than 25 students – five colleges; 25 to 50 students – four colleges; more than 50 students – five colleges.

b) The reported numbers of redirected students over the span of time were: less than 25 students – five colleges; 25 to 50 students —three colleges; more than 50 students — five colleges.

c) The reported number of students screened out were: less than 25 students – five colleges; 25 to 50 students — none; more than 50 students – three colleges. There seemed to be especially great reluctance to disclose figures of those screened out. It should be kept in mind that the numbers of students affected probably depends upon the size of the college, as well as upon other factors. It does seem to be true, however, that the number of students who are affected by these programs is small, without regard to the size of the college.

Inquiry was made as to whether the respondents considered their programs to be satisfactory in terms of the effect upon the quality of later student teaching and professional attitudes. Sixteen said yes, but 14 of these, while favoring the program, were reserved in their praise. Four of these described the results as "imperfect but helpful." One referred to the extreme difficulty of full confidence in the evaluations, and one stated that even though his college had such a program, some "poor risks" still slip past the screen.

Several difficulties or shortcomings have shown themselves in the application of the redirecting and screening procedures. Four respondents wrote of the need for more exact and precise ways of measuring the traits. Two others mentioned the inability of the members of the staff or committee to agree as to the disposition of an individual case. One stated that staff members' sympathy leads to softness, which on occasion defeats the screen. Another mentioned the reluctance of faculty members to accept this or other new programs, with resultant lack of cooperation and weakening of the structure. One opinion was that each student majoring in a teaching field should undergo a psychiatric examination, and that this would lend strength and teeth to final decisions. Three bemoaned the heavy load of faculty members as reducing the effectiveness of the program.

Summary and Conclusions

The returns show that variety is the word for ways of evaluating and doing something about students' poor teaching personalities. At one leading university in the East, all teaching majors are required to take the sophomore professional experience course, which is designed specifically as a screening device or hurdle, to be met before proceeding to other work in preparation for teaching. At a few other colleges, no formal program exists for the purpose under consideration. One respondent indicated that no such evaluative program was in effect until the junior year at his college.

In spite of the variety of the variety of means of evaluating teaching personality of students, the study brought out the broad area of agreement among those with responsibility for preparing good teachers. An impressive majority of respondents stated that they did have a procedure for identifying and redirecting education majors with teaching personality problems. In more than one-half of the reporting colleges, the procedure was formal and written, in addition to the informal, non-channeled appraisal of students' performances. Near unanimity was noted in that the evaluating was done or was initiated by the instructor of the professional experience course, either freshman or sophomore. The decision as to the final disposition of a case under study most commonly came from a committee. The responsibilities for such decisions are grave, and one would expect that they should call for much investigation, discussion, and cooperative agreement among the several members of the committee.

The number and proportion of students screened out or redirected would seem to be small. Despite this, 14 out of the 16 respondents said that the program was effective and beneficial in terms of improved professional attitudes and of the quality of later student teaching. Most of those favoring their own personality evaluation program, however, regretted the lack of perfection there.

Such programs admittedly are not so objective and neat as measurements based upon indisputable facts, data, and figures as found in standardized testing programs. Appraisals of teaching qualities must be subjective, and they demand the placing of confidence in the judgment of those responsible for the preparing of good teachers. A writer in the 1960 yearbook of the Association of Student Teaching pointed out that the supervisors' judgment and subjective interpretation of qualities possessed by student teachers should be a part of the evaluation of the success of student teachers.* The fact that such qualities are not weighable and objective does not relieve the burden of those with responsibility for providing competent teachers for the children

---

*Leander L. Boykin. "Principles of Evaluating in Student Teaching," Association for Student Teaching, 39th Yearbook, (Cedar Falls, Iowa, 1960), pp. 20-21.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 7, March 1963)

Teacher Enthusiasm*

VICTOR E. MASTIN
Meredith Junior High School
Des Moines, Iowa

THE CENTRAL PURPOSE of this study was to ascertain whether boys and girls are likely to learn more under teachers who appear to be enthusiastic about the topics, materials, and ideas which they present than under teachers who are indifferent toward such topics, materials, and ideas.

With the expectation of obtaining further help in understanding this problem, a study was also made of the degree to which pupil attitudes are affected by the attitudes of their teachers, and the relationships between intelligence and reactions to the contrasted attitudes and presentations.

To obtain data, an experiment was planned which would provide real teaching situations wherein the attempt might be made to evaluate those influences as they would make themselves felt in normally organized classroom groups working in their home environments.

The plan of the experiment called for a teacher to present an illustrated lesson to his class in such a manner as to convey to the group a feeling that he had an indifferent attitude toward the ideas and pictures being presented, and toward the subject of the lesson. The experimenter would then administer an objective test to determine what facts had been learned, and an attitude scale to ascertain how the pupils felt about the lesson subject, the teacher's work, and the lesson itself, and how much interest they had in learning more. Approximately one week later, with the environmental, time, and motivational conditions as nearly identical as possible, the same teacher would present a second lesson in such a manner as to convey to his pupils the impression that he was enthusiastic about the ideas and illustrative materials of the lesson, and the subject covered by the lesson. The experimenter would then give the test over that lesson and administer the same attitude scale.

In order to arrive at a satisfactory description of the actions and possible appearance of a teacher who might impress his pupils as having an attitude of indifference toward, or enthusiasm for, the lesson materials and ideas, and the subject being presented, contrasting lists were prepared which purported to describe the means whereby such attitudes might be made apparent to the pupils of the class. The list was evaluated by a board of experts, then further revised for use by teachers, and by the experimenter.

For teaching purposes, two illustrated-lecture type lessons were prepared, one on ancient Egypt, and one on ancient Rome and Pompeii. The illustrative materials consisted of a set of fifteen, 11- by 14-inch mounted photographs for each lesson, and a map of the Mediterranean area.

A 102 question, four-item, multiple-choice test was prepared for each lesson. A 20-statement, five-step "Strongly Agree" to "Strongly Disagree" attitudes scale was also constructed and attached to each of these tests. The statements, randomly arranged, included five in regard to each attitude area. The Otis Quick Scoring Mental Ability Test was used to obtain information about the intelligence of these pupils.

Since it was thought that the pupils' reading ability might influence their scores on a test of the length used for this experiment, the questions and all the possible answers were read aloud to them, as were the statements and options of the attitude scale.

For the main experiment, 20 classes were used, seven having been sixth-grade groups from as many schools, and 13 having been seventh-grade sections from three junior high schools. The classes and schools were selected as being representative of the city's sixth- and seventh-grade population and were chosen, a) on the basis of representativeness as to geographical location and socio-economic factors, and b) on the condition that the school's entire grade would be available without splitting classes. Five hundred sixty-one members from these classes were present for both lessons. The comparisons made and the conclusions drawn are based on the data from these 20 classes and 561 pupils.

---

* This report is of a Doctoral Study directed by Professor Hubert H. Mills, University of Colorado.

The selection of teachers was left to chance in that the classes and schools were first chosen, and then the teachers who had these classes for core or major subjects were invited to participate. Since no factors other than chance were known to have been involved in the assignment of teachers to these classes, it was believed that this approach would result in the enlistment of persons who would be representative of the types of personalities met by public school pupils. This belief seemed to have been borne out by evidence from the observation of these teachers at work during the two presentations, and from the pre-planning and post-presentation discussions.

There were four possible orders of presentation for the two lessons and two types of presentation. Twenty classes made it possible to repeat each series five times.

Upon examination, the data revealed that the class mean for the lesson which had been taught with apparent enthusiasm, whether that lesson was presented first or second, and whether it was the Roman or Egyptian story, was higher for 19 of the 20 classes than the mean of the lesson presented with indifference. The difference in means of these 19 classes ranged from .13 to 14.68 points. In 15 of these classes, the difference was great enough to be statistically significant at the one percent level. For one class, the difference in means favored the indifferently presented lesson. This difference, of 2.15 points, was such that it was significant at the five percent level.

A study of the individual scores showed that 435, or 77.54 percent of 561 pupils, made larger scores on the test over the enthusiastically presented lesson, 102, or 18.18 percent made larger scores on the test following the indifferently taught lesson, and 24, or 4.28 percent made the same on both.

Four hundred fifty-five pupils were available for the Otis test. The product-moment correlations between the Otis and indifferent scores ranged from .23 to .70, and between the Otis and enthusiastic scores from .29 to .69. The Otis to difference correlations were comparatively small, ranging from -.36 to .41, with nine negative and seven positive. These data were interpreted to indicate that this experiment furnished no evidence of significant relationships between individual intelligence, as measured, and reactions to the contrasted presentations.

Three hundred eighty-one, or 67.91 percent of the pupils gave the enthusiastic teacher higher scores on the statements of the teacher section of the attitude scale, while 119, or less than one-third as many, gave higher scores to the indifferent teacher. In no class did a majority of the pupils rate the indifferent higher than the enthusiastic teacher.

When all the section scores were combined from one scale and compared to the total score on the other scale, it was found that 381, or 67.91 percent of the pupils, had given higher total scores to the statements on the attitude scale marked following the enthusiastic presentation, while 162 had given the higher total score to the scale marked following the indifferent presentation. A majority of the pupils in 16 of the 20 classes gave higher score totals to the attitude scale marked following the enthusiastically presented lesson.

When the total numbers of pupils who gave higher scores on the different sections of the attitude scale to the enthusiastic teacher, to the enthusiastically presented lesson and people, and to their interest in learning more about these people, were compared by the chi-square technique to the numbers which would have been expected under the null hypothesis, significant levels above one percent were found for all sections, and for the net totals as well.

The results of this experiment would seem to indicate that the attitudes which teachers appear to have toward the topic, materials, and ideas which they present to their classes do influence the factual learning of the pupils in their classes, and the attitudes of those pupils toward the topic, materials, and ideas.

Since the findings indicate that enthusiasm on the part of teachers has a real influence on pupil learning and on pupil attitudes, it becomes important, (a) that steps be taken to assure, in presently organized classrooms, that the functioning of this element becomes better understood and its power more often used, (b) that this element be given its proper place and weight in the minds of those who are responsible for the selection and training of teachers, and (c) that this force be recognized and accounted for as a potentially powerful influence which is often, without recognition, at work in experimental research on educational problems.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 7, March 1963)

The Comprehension Subtest of the Wechsler Adult Intelligence Scale as an Indicator of Social Awareness

ROBERT M. ROTH
Hampton Institute

CLINICIANS CAN interpret WAIS protocols on two levels. The first level relates to intellectual functioning. The second level relates to personality evaluation. According to Wechsler's assumptions (2), the responses to each subtest can be inpreted both as an expression of intellectual functioning and as an expression of personality dynamics.

Wechsler assumes that the Comprehension subtest "... almost always tells of something about the subject's social and cultural background. Offhand it might be termed a test of common sense, and it is so-called on the Army Alpha. Success on the test seemingly depends on the possession of a certain amount of practical information and a general ability to evaluate past experience."

The purpose of this study was to test Wechsler's assumptions. "Common sense" and "social and cultural backgrounds" are assumed to be expressed in terms of social awareness in this study. This concept refers to one's understanding his world and his role in it. It is further assumed that those who indicate social awareness, function effectively in social situations, and that their social effectiveness is perceived by their peers.

Bonney (1) has reported findings which seem to support this contention. He reports that in a men's college dormitory, high sociometric choice men frequently showed sincere and objective interest in others, consideration for others, strongly positive self-regarding attitudes, and active participation in more mature and socially approved group activities.

This study has attempted to test the hypothesis that those considered to be more socially aware by their peers will score higher on the Comprehension subtests than those considered to be less socially aware by their peers.

Procedure

In a women's residence hall at Hampton Institute, freshmen students were grouped by corridors and other natural groups within the buildings. Five groups were chosen. The women in each group had lived in the group for a minimum of 8 months. Each group had consisted of 19 members. Each member was asked to rank the women in her group according to friendliness, well-roundedness, and likeable personality. These factors were selected by the students administering the study as the element of social awareness most likely to be identified by the subjects, as a basis for sociometric choice. In terms of total scores, those students who fell into the top and bottom three ranks of each group were identified as more and less socially aware, respectively. From these rankings, 30 students were selected for the experimental group. The Comprehension subtest of the WAIS was administered to these selected subjects. Fifteen in the top and 15 in the bottom rankings, six from each group. Raw scores were converted into scaled scores. The data were tested with t-test and F-ratio analyses. The tests were administered and the findings analyzed by the students in an advanced Psychology class.

Results

No significant differences were found for the scores on the Comprehension subtest of the WAIS between those subjects considered most socially aware and least socially aware by their peers. These results are indicated in Table I. The subjects' mean scores did not differ significantly. The assumption that the top and bottom ratings represented essentially different populations did not hold when subjected to the F test.

Discussion and Conclusions

Although it has been shown that the interpretations of test by clinicians may be valid, even where the instrument has not adequately been validated otherwise, it is important that assumptions about the meaning of responses to test items be tested. This is particularly true of the WAIS, of which widespread use has been made in clinical interpretations.

The purpose of this study was to test the assumptions that the Comprehension subtest of the WAIS measured differences in "social awareness". Subjects were selected for social awareness on the basis of peer group sociometric choices. No significant differences were found in WAIS scores on the Comprehension subtest between students receiving high and low sociometric rankings.

TABLE I

DIFFERENCES IN COMPREHENSION SUBTEST SCORES BETWEEN MORE AND LESS SOCIALLY AWARE SUBJECTS

	Mean	S.D.	t	P	F	P
More	13.27	3.25	.78	NS	1.05	NS
Less	12.40	3.07				

On the basis of this finding, it seems that two possible conclusions are tenable. The first involves the invalidity of Wechsler's assumptions regarding the social awareness aspect of the Comprehension subtest. Further testing of the assumption regarding the non-intellectual factors of the other subtests is indicated. Secondly, the test might be measuring aspects of intellectual functioning which are not expressed in behavior. Conceptions regarding the dissociation between internal and external expressions of intellectual function need to be developed.

REFERENCES

1. Bonney, M. E., and Dryer, A. H. "A Study of Some Factors Related to Sociometric Status in a Men's Dormitory," Sociometry, XVI (1953), pp. 287-301.
2. Wechsler, D., The Measurement of Adult Intelligence, Third Edition, (Baltimore, The Williams and Wilkins Company, 1944).

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 7, March 1963)

Personality Needs of Academically Superior Students and Their Parents*

KURT HAAS
State University College
New Paltz, New York

ACADEMICALLY SUPERIOR college students, that is, those that obtain high grade-point averages, have been reported to possess a number of different needs and drives which distinguish them from average undergraduates (1, 2, 5). In addition, the personality needs of the parents of the successful students have been held to be distinct from those of other adults (4). It was the purpose of this study specifically to examine some of the personality attributes of the academically superior student and those of his parents.

Procedure

A group of 25 upper-class undergraduate students in a Midwestern state university with a grade-point average of 3.60 or better (where 4.00 equals "A") were administered the Edwards Personal Preference Schedule (EPPS). The EPPS was also completed, by mail, by the 25 mothers and 25 fathers of these students.

The EPPS provides measures of 15 personality variables. These are as follows:

1. Achievement
2. Deference
3. Order
4. Exhibition
5. Autonomy
6. Affiliation
7. Intraception
8. Succorance
9. Dominance
10. Abasement
11. Nurturance
12. Change
13. Endurance
14. Heterosexuality
15. Aggression

Normative data are given by Edwards (3) for college students and for adults, so that comparisons of specific scores will reveal whether those scores are within limits or markedly high or low. Accordingly, the means of the 15 personality needs computed for the successful students and their parents were compared with the norms provided by Edwards. These comparisons were made by using *z* scores. The mean score on each personality variable for each group was also converted into a percentile in order that the relative position for each score in the sample could be contrasted with the normative distribution.

Results

None of the 15 personality needs measured by the EPPS was found significantly different ($p \neq .05$) from the norms. The percentiles corresponding to the mean scores are reported in Table I in order to demonstrate the relative standing of the students and parents. With one exception (deference for male students) none of the variables meets Edwards' (3) criteria for being very high or low in departure from average.

Discussion

It appeared as if the academically superior student and his parents could not be said to differ significantly from appropriate normative populations in terms of the personality needs assessed on the EPPS. Contrary to common supposition, the students did not evidence higher achievement needs or significant drives for order, endurance, etc. Similarly, the parents were not found to be significantly characterized by dominance or some of the other personality variables frequently assumed.

A number of explanations might serve to account

---

* The author is indebted to the Committee for Scholarly Advancement, Bowling Green State University, for a grant in aid of this research.

TABLE I

EPPS PERCENTILE SCORES FOR STUDENTS AND THEIR PARENTS

	College Students		Parents	
Variable	Males	Females	Fathers	Mothers
Achievement	58	72	56	50
Deference	96	62	52	66
Order	84	72	40	56
Exhibition	37	42	57	42
Autonomy	52	53	62	54
Affiliation	28	50	58	35
Intraception	64	28	61	52
Succorance	42	67	42	46
Dominance	21	59	69	56
Abasement	75	66	56	51
Nurturance	46	41	47	31
Change	33	28	46	60
Endurance	75	76	44	64
Heterosexuality	45	45	50	66
Aggression	57	60	62	71
Consistency Score	86	63	70	66
N	(12)	(13)	(25)	(25)

for the failure to obtain significant differences on the EPPS. It is possible that the lack of significant results is due to failures in design or weaknesses within the EPPS. It appears, however, as if the most likely explanation for the lack of uniqueness for the scores of the parents and the students may be simply, that successful students are successful not because of distinctly unusual personality needs but because they have more ability and apply it more diligently.

Summary

The personality needs, as measured by the EPPS, of academically superior college students and of their parents were compared to the norms reported for a general college sample and a general adult sample, respectively. Significant differences were not obtained. A number of explanations were suggested, among which it was held likely that superior students were probably not very different from other college students except in terms of ability and diligence.

REFERENCES

1. Bendig, A. W. "Predictive and Postdictive Validity of Need Achievement Measures," Journal of Educational Research, LII (1958), pp. 119-120.
2. Horrall, B. M. "Academic Performance and Personality Adjustment of Highly Intelligent College Students," Genetic Psychological Monographs, LV (1957), pp. 3-83.
3. Edwards, A. L. Manual for the Edwards Personal Preference Schedule (New York; Psychological Corporation, 1954).
4. Weigand, G. "Adaptiveness and the Role of Parents in Academic Success," Personnel Guidance Journal, XXXV (1957), pp. 518-522.
5. Zeaman, J. B. "Some of the Personality Attributes Related to Achievement in College: A Comparison of Men and Women Students," Doctor's Thesis, Michigan State University, Dissertation Abstracts, XVIII (1958), pp. 290-291.

Among Those Present

CLIFTON M. CLAYE, Arkansas EdD, has served as teacher and principal for 14 years in Arkansas schools, and is now assistant professor of secondary education at Texas Southern University, Houston.

KURT HAAS, assistant professor of psychology at Bowling Green State University, holds a PhD from Penn State, where he was a counselor. He has appeared in *The Journal of Psychology* and is presently a consultant to an Ohio county school system.

KENNETH H. HOOVER, Washington EdD, is associate professor of education at Arizona State with teaching experience in the high schools of Washington, at Montana State, and San Francisco State. He is co-author of *Scientific Foundations of Education.*

VICTOR E. MASTIN holds an EdD from Colorado. He is principal of James Callanan Junior High in Des Moines, Iowa. He has taught at Michigan State, Colorado, and Drake.

LEBERN N. MILLER, Texas EdD, supervises student teaching at Trenton State College, Trenton, N. J. He has taught in Illinois, California, Minnesota, and Texas.

LEON D. RADAKER, psychologist with the Butler, Pennsylvania, Area Schools, holds degrees from Slippery Rock, Pittsburgh, and Penn State.

J. WESLEY SCHNEYER holds degrees from West Chester State College, Pennsylvania, and the University of Pennsylvania. He is presently assistant professor of education, and supervisor of the high school, college, and adult reading programs of the Reading Clinic at the University of Pennsylvania.

MILDRED C. ROBECK taught in the schools of Wisconsin, Illinois, and Washington for 15 years before becoming assistant professor of education and director of The Reading Clinic for Children at the University of California, Santa Barbara. Her three degrees are from the University of Washington.

BERTHA BOYA THOMPSON holds degrees from Slippery Rock State College, Pennsylvania, Miami University, Ohio, and Indiana University. She has had 12 years of public school teaching and two in directing a private kindergarten. She is presently coordinator of social studies at the Talawanda High School, Oxford, Ohio, and teaches evening courses for Indiana University at the Earlham College — Indiana University Education Center at Richmond, Indiana. Her husband is professor of geography at Miami.

WARREN J. VORREYER is dean of the college and assistant to the president at Southern State Teachers College, Springfield, South Dakota. His EdD is from the University of Wyoming. He has held vocational education posts in Illinois, California, Wyoming, and South Dakota.

WILLIAM V. WAGNER is a West Virginia native who holds an EdD from his state university. Presently assistant professor of guidance at West Virginia University, he has taught in West Virginia public schools and was assistant registrar at Marshall University, Huntington, West Virginia.

MOVE, MOVING, MOVED? Won't you help us keep your Journal coming to you on time by keeping us posted on your whereabouts?

Your Name ________________________________

Old Address ________________________________

____

New Address ________________________________

____

DEMBAR EDUCATIONAL RESEARCH SERVICES, Inc.
Post Office Box 1148, Madison 1, Wis.

Executive Editors

WILSON B. THIEDE, Professor of Education, Educational Administration, and Agricultural and Extension Education, The University of Wisconsin, Madison, Wisconsin. **Chairman.**
THEODORE HARRIS, Professor of Educational Psychology and Education, The University of Wisconsin, Madison, Wisconsin.
L. JOSEPH LINS, Professor and Coordinator of Institutional Studies, The University of Wisconsin, Madison, Wisconsin.
PHILIP LAMBERT, Professor of Educational Psychology and Educational Administration, The University of Wisconsin, Madison, Wisconsin.
JOHN SCHMID, Professor of Education and Chairman of the Department of Educational Foundations, University of Arkansas, Fayetteville, Arkansas.

Editorial Board

JACOB O. BACH, Head, Educational Research Bureau, Southern Illinois University, Carbondale, Illinois.
WILLIAM A. BROWNELL, Dean, School of Education, University of California, Berkeley, California.
ROBERT A. DAVIS, Professor of Educational Research, George Peabody College for Teachers, Nashville, Tennessee.
GUY D. McGRATH, Dean, College of Education, Arizona State University, Tempe, Arizona.
HAROLD E. MITZEL, Director, Office of Research and Evaluation, Division of Teacher Education, College of the City of New York, New York City.
WILLARD C. OLSON, Dean, College of Education, University of Michigan, Ann Arbor, Michigan.
ARTHUR E. TRAXLER, Executive Director, Educational Records Bureau, New York 32, New York.

Associate Editors

HAROLD M. ANDERSON, Associate Professor of Education, University of Colorado, Boulder, Colorado.
VERNON E. ANDERSON, Dean, College of Education, University of Maryland, College Park, Maryland.
RUDYARD, K. BENT, Professor of Education, University of Arkansas, Fayetteville Arkansas.
EMMETT ALBERT BETTS, Research Professor, University of Miami, Coral Gables, Florida.
LAWRENCE P. BLUM, Assistant Director, Student Personnel Services, University of Wisconsin-Milwaukee, Milwaukee 3, Wisconsin.
J. MURRAY LEE, Chairman, Department of Elementary Education, Southern Illinois University, Carbondale, Ill.
D. WELTY LEFEVER, Professor of Education, University of Southern Caifornia, Los Angeles, California.
S. E. TORSTEN LUND, Associate Professor of Education, University of California, Berkeley, California.
WILLIAM A. McCALL, Lecturer in Education, University of Miami Coral Gables, Florida.
DELYTE W. MORRIS, President, Southern Illinois University, Carbondale, Illinois.
VALWORTH PLUMB, Chairman, Division of Education and Psychology, University of Minnesota (Duluth Branch), Duluth, Minnesota.
DAVID G. RYANS, Rand Corporation, Santa Monica, California.
LESTER B. SANDS, Professor of Education, University of California, Santa Barbara, Goleta, California.
MANDEL SHERMAN, President, and Consultant, Emotional Health Association Los Angeles, California.
MILDRED B. STANTON, Consultant, Special Education, State Department of Education, Hartford 15, Conn.
W. W. THEISEN, Professor of Education, Marquette University, Miwaukee, Wisconsin.
M. R. TRABUE, Professor of Higher Education, University of Kentucky, Lexington, Kentucky.
J. WAYNE WRIGHTSTONE, Director, Bureau of Educational Research, Board of Education of City of New York, Brooklyn, New York.

Editors Emeriti

E. J. ASHBAUGH, Dean Emeritus, School of Education, Miami University, Oxford MR50, Ohio.
HARRY J. BAKER, 19050 Wiltshire, Lathrup Village, Michigan.
FOWLER D. BROOKS, Director Emeritus, Department of Psychology, DePauw University, Greencastle, Ind.
LEO J. BRUECKNER, Professor Emeritus, University of Minnesota, Minneapolis, Minnesota.
B. R. BUCKINGHAM, 33 Chatham Circle, Wellesley Hills, Massachusetts.
EDGAR A. DOLL, Director Emeritus, Devereux Schools, Devon, Pennsylvania.
HARL R. DOUGLASS, Director Emeritus, School of Education, University of Colorado, Boulder, Colorado.
E. DUNCAN GRIZZELL, Dean Emeritus, University of Pennsylvania, 640 Maxwelton Court, Lexington, Ky.
ERNEST HORN, Professor Emeritus of Education, State University of Iowa, W307 East Hall, Iowa City, Iowa.
EARL HUDELSON, Professor Emeritus of Education, West Virginia University, Morgantown, West Virginia.
A. M. JORDAN, Professor Emeritus, University of North Carolina, Chapel Hill, North Carolina.
ALFRED S. LEWERENZ, Director Emeritus of Evaluation and Research, 4107 Holly Knoll Drive, Los Angeles, California.
A. R. MEAD, Director Emeritus of Educational Research, University of Florida, 1719 N. W. 6th Ave., Gainesville, Florida.
M. J. NELSON, Dean Emeritus, Iowa State Teachers College, 6033 West Orange Drive, Glendale, Arizona.
JESSE B. SEARS, Professor Emeritus, Stanford University, Stanford, California.
CARLETON W. W

The Journal of Educational Research

Dedicated to the Scientific Study of Education

VOLUME 56 **APRIL 1963** **NUMBER 8**

It Says in the JER ---------- 394

Can the School Curriculum Make a Frontal Approach to Mental Health?—*Gail M. Inlow* ---------- 395

Changes in Self-Awareness During the High School Years: A Study of Mental Health Using Paper-and-Pencil Tests—*John E. Nickols, Jr.* ---------- 403

An Analysis of the Quality of the Behaviors of Principals as Observed and Reported in Six Critical Incident Studies —*Bernice Cooper* ---------- 410

Academic Achievement and Adjustment of Underage and Overage Third Graders—*Marian L. Carroll* ---------- 415

A Comparison of Two Methods for Determining Difficulty in a Multiple Choice Test—*Margery V. Hawes* and *Gerald C. Helmstadter* ---------- 420

Educators' Attitudes Toward Educational Practices—*Robert S. Miller* ---------- 424

Ability of Poor Readers to Discriminate Paired Associates Under Differing Conditions of Confirmation—*Wayne Otto* 428

Readers Who Lacked Word Analysis Skills: A Group Diagnosis—*Mildred C. Robeck* ---------- 432

How Intermediate Children Rank the Subjects—*Wayne L. Herman, Jr.* ---------- 435

An Investigation of the Attitudes of Parents of Unsuccessful and Successful Readers—*Dorothy P. Macdonald* ---------- 437

Age, I.Q., and Improvement of Reading—*Norman M. Chansky*

The Relationship Between Pupil Self-Evaluation, Teacher Evaluation of the Pupil and Scholastic Achievement—*Ira J. Gordon* and *Patricia C. Wood* ---------- 441

Field News ---------- 444

Among Those Present and Readers Write ---------- 447

The Journal of Educational Research is published at Madison, Wisconsin, monthly 10 times a year (the May-June and July-August issues being combined). The subscription price is $7.50 a year, plus $1 postage for all subscriptions outside the continental United States. Single copies, $1. Second class postage paid at Madison, Wisconsin. Copyright 1963 by Dembar Educational Research Services, Inc. Address all correspondence care of DERS, Box 1605, Madison 1, Wisconsin.

Published by **DEMBAR EDUCATIONAL RESEARCH SERVICES, Inc.** WALTER FRAUTSCHI, President. WILSON B. THIEDE, Vice President and Publisher. Prof. CLARENCE A. SCHOENFELD, Assistant

It Says in The JER...

CHANSKY FINDS NO EMPIRICAL SUPPORT for the belief that children with high IQ's make the greatest progress in remedial reading. He suggests the need to seek another criterion for making selections for classes in remedial reading.

AN EXPLORATORY STUDY by Macdonald suggests interesting implications in the hypothesis that a child's relation with his parents may affect his learning to read.

THE SOCIAL STUDIES RANK low in interest among intermediate children, Herman finds. Spelling, arithmetic, and English rank high. Herman suggests the "expanding environment" concept of teaching social studies could be responsible for the low interest.

DO EDUCATORS AGREE with the current lay critics of educational philosophy and practices? No, says Miller after a study of professional attitudes.

AFTER COMPARING two methods for determining item difficulty in a multiple-choice test, Hawes and Helmstadter conclude it makes little difference whether you determine the percentage of examinees who correctly answer a question or whether you count the number of responses required by an individual to obtain the correct answer.

HIGH SCHOOL SENIORS achieve significantly higher scores on a paper-and-pencil mental health test than do freshmen, Nickols reports. The data also suggest that intelligence is positively related to score gains made during the three-year period.

Continued on page 402

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 8, April 1963)

Can the School Curriculum Make A Frontal Approach to Mental Health?

GAIL M. INLOW
Northwestern University

THE UNDERLYING purpose of this research from its inception was to discover if psychiatrists and clinical psychologists believe that elementary and secondary school personnel can, with profit, incorporate mental health concepts directly into the curriculum.

It was accepted as fact that many children and adults are unable to measure up to acceptable standards of emotional maturity and that their failures lead to individual and societal disorders which, in turn, make for unnecessary tensions and resulting unhappiness. The merits of therapy after the fact of illness have been demonstrated too often to be questioned. Unfortunately, however, therapy is not accessible to all who need it; and, by definition, it is a sequel to illness, not a means of prevention.

Therefore, what seems to be needed, even though elusive of attainment, is a prophylactic approach to emotional disorders. Specifically, in the school milieu, is it possible that selected mental health concepts and understandings can be made a part of the elementary and secondary school curriculum? And, if a part, what concepts and understandings should be included? Also, how should they be treated in terms of teaching method?

The problem may be restated as follows: Will the teaching about mental-health concepts at the appropriate readiness level act as a preventive or alleviator of emotional disorders at a later time?

The Procedure

The sample. It was hypothesized that the best source of information about mental-health practices would be individuals who make a profession of mental health. To narrow the field, the investigator decided to include only psychiatrists and clinical psychologists in the study.

First of all, he made a list of the psychiatrists whose names appear in a designated place on the various pages of the Biographical Directory of Fellows and Members of the American Psychiatric Association, 1958. This list represented approximately one-twentieth of all the names in the Directory. Next, from the American Psychological Association Directory, 1958, he made a list of the names of every twentieth clinical psychologist, fellow or member. The respective lists included the names of 444 psychiatrists and 101 clinical psychologists, a total of 545 individuals.

In an effort to pre-determine what proportion of the potential sample would be willing to participate in the anticipated study, he selected at random the names of 54, or ten percent, of the potential sample members--44 psychiatrists and ten clinical psychologists. A post card was subsequently mailed to these individuals whereon they were asked if they would be willing to complete a questionnaire on mental health designed to assist public and private school personnel in their efforts to reduce the extent of emotional disorders among children. Four of the cards sent to psychiatrists were returned indicating changes of locations without accompanying forwarding addresses. Of the 50 remaining persons, only six of the forty psychiatrists, or 15 percent, and three of the ten clinical psychologists, or 30 percent, expressed a willingness to cooperate in the study.

Despite the forewarning of a low percentage of returns, the investigator decided to proceed with his research plans. He subsequently mailed a questionaire to the remaining 400 psychiatrists and 91 clinical psychologists. Changes of address of 38 reduced the sample to 373 psychiatrists and 80 clinical psychologists, or a total of 453 sample members.

Returns. Replies were received from 58, or 15.5 of the psychiatrists, and 25, or 32 percent of the clinical psychologists. The combined percent of returns was 18.3. One of the replies from a psychiatrist was not usable, leaving 82 participants --57 psychiatrists and 25 clinical psychologists.

The investigator makes no attempt to establish the representativeness of the sample. He will only report and make an effort to interpret the data that the 82 individuals submitted to him.

The questionnaire. In view of the fact that the sample members resided in all parts of the United States and many countries abroad, a questionnaire was selected as the only logical data-gathering medium. It was mailed on May 14, 1960. The last reply, the 83rd, was received on July 28. The "heart" of it was centered in the following three questions:

a. Will you identify in the field of personality dynamics the five to ten concepts and understandings which you believe are the most meaningful for inclusion in an elementary or secondary-school curriculum

b. Will you next indicate in the designated column the earliest age when the children or youth could profit from a learning experience which emanates from the suggested concepts?

c. Whenever you believe that a specific teaching method is indicated by the concept, will you identify the method?

The last part made an attempt to secure information of significance to teacher-education institutions.

> Assuming that the teacher of the afore-mentioned concepts is a well-adjusted, emotionally mature individual, what specialized training in the field of personality dynamics do you believe would be a minimum prerequisite for his teaching role?

The questionnaire was structured as little as possible in an effort to motivate the optimum of spontaneity.

The Findings

Data contained in the 82 meaningful questionnaire returns will be reported under the following categories--professional information about the sample, curriculum-content suggestions, curriculum-methods suggestions, and teacher-education suggestions.

Professional information about the sample. The 25 clinical psychologists had earned at some previous time the Ph. D. Degree, the 57 psychiatrists, the M. D. Degree. Of the 82 returns, 68 were categorized by the investigator as complete--complete in the sense that they reflected a well thought out point of view and presented meaningful information in support of the point of view. The remaining 14 were categorized as fragmentary. They provided significant information, for the most part, but within narrowly conceived limits.

In reply to the question "What is your professional orientation?" 25 indicated that they were psychoanalysts; 13, neuro-psychiatrists; 21, psychiatrists or general psychiatrists; three, child psychiatrists; two, medical administrators; and 25, clinical psychologists. The fact that the above figures total 89 instead of 82 reflects the fact that a number of the sample participants had dual professional roles or supplied two titles to describe their roles more clearly.

Seventy-five reported that they were practicing therapy, four that they were not, and three did not answer the item. Of the 75 who were actively engaged in the practice of therapy, 38 worked only with adults, 37 worked only with youth or with both youth and adults. The mean years in therapy of the sample group were 11.6, the median, 11.1.

Their attitude toward the research project may be categorized essentially as follows: Forty-eight were positive, i.e., their suggestions were made without reservations and they seemed optimistic about the merits of teaching mental-health concepts directly; ten were ambivalent--they provided meaningful information but were agnostic or mildly skeptical about the research project if implemented in practice; 23 were negative--five because they believed that too much preoccupation with self is harmful, six because an intellectual approach to mental-health concepts "will not change people emotionally", and 12 because of miscellaneous reasons other than the two stated. One answer was too vague to be classified. Of the replies from the 57 psychiatrists, 28 were positive, nine ambivalent, 19 negative, and one too vague to classify; of the replies from the 25 clinical psychologists, 20 were positive, one ambivalent, and four were negative. It may be observed, therefore, that the clinical psychologists were somewhat more optimistic about the proposed approach than were the psychiatrists.

Curriculum-content suggestions. In Table I are listed the psychological concepts of personality, some general, some specific, which the respondents proposed as practicable for inclusion in a school curriculum. The concepts are divided according to the type of specialist who proposed them: 1) psychoanalyst, 2) neuro-psychiatrist, psychiatrist (listed as such), or general psychiatrist, and 3) clinical psychologist. Category No. 2, above, will hereafter be referred to by the term non-psychoanalytical psychiatrist to differentiate it from the psychoanalytical school. Also contained in the table is the median age when the sample members believed the concepts should first be introduced into the curriculum.

The captions of the grouped concepts are those of the investigator. The original suggestions were received, listed individually on cards, and then categorized. To make the grouped captions more meaningful, word-for-word concepts as presented by selected sample members have been included following each categorized listing. An inevitable amount of overlapping admittedly exists between and among the various concepts and groupings.

Table I is to be interpreted as follows. For instance, in Item 1, 36 participants (ten of the 25 psychoanalysts, ten of the 32 non-psychoanalytical psychiatrists, and 16 of the 25 clinical psychologists) proposed that one or more of the defense mechanisms could, with profit, become a part of the school curriculum. The median age for the mechanism(s) to be introduced was 10.8 years.

The most prolific contributors of mental health proposals were the clinical psychologists. Twenty-five of them made a total of 90 suggestions in contrast to 80 made by the 25 psychoanalysts and 59 by the 32 non-psychoanalytical psychiatrists.

Although the suggestions made by the three groups, for the most part, were comparable in emphasis, differences of some significance may be noted. The clinical psychologists placed their greatest emphasis on the defense mechanisms, Item 1 (16 of 25, or 64 percent), on the emotions, Item 2 (14, or 56 percent), on the symptomatic reactions, Item 3 (11, or 44 percent), and on inter-personal relationships, Item 4 (ten, or 40 percent).

The proposals by the psychoanalysts, in rank order of priority, were sex manifestations, Item 6, and defense mechanisms, Item 1 (ten, or 40 percent each), and the concept of self, Item 5, along with the emotions and drives, Item 2 (9, or 36 percent each).

The non-analytically oriented psychiatrists placed in top priority the defense mechanisms, Item 1 (ten, or 31 percent), the emotions and drives, Item 2 (7, or 22 percent), with the topographical, Item 7, and inter-personal relationships, Item 4, tied for fourth place with six supporting suggestions, or 18 percent of the total possible.

On a comparative basis, the psychoanalysts made significantly more proposals than the other two groups only on the factor of sex, Item 6. The non-analytically oriented psychiatrists led each of the other two groups in only one area, the topographical concepts, Item 7, but by only a small ratio of 6-5-3. The clinical psychologists led the field on a comparative basis in the areas of the defense mechanisms, Item I; the emotions, Item 2; symptomatic reactions, Item 3; interpersonal relationships, Item 4; and general concepts of mental health and illness, Item 9.

Of interest in the replies is the age level children to whom the participants believed the concepts could first be introduced. The median age suggestions range from 9.0 years for the area of interpersonal relationships to 14.5 years for the concept of multiplicity and complexity of factors which cause poor mental health. In general, the range is from the mid-latency period to the period of early adolescence. However, it should be remembered that the total range for each of the concepts is much more extensive than that of the medians. Several participants proposed a range extending "from birth" for one concept to "over 18" for another.

In an effort to determine what factors other than professional orientation might have been associated with differences in responses, the investigator divided the sample group on the basis of years in therapy and age-level of clients.

On the factor of years in therapy, he compared the responses of psychiatrists, all types, who had been therapists for fewer than nine years, with those of their counterparts who had been therapists for 21 years or longer. In each group were 14 individuals. When the responses of the two groups to the various items were compared, it was discovered that only on the item of interpersonal relationships was there any appreciable difference. Four of the psychiatrists who had practiced 21 years or longer, but none who had practiced fewer than nine years, included the item in their curriculum suggestions. Because of the small size of the groups, however, significant differences were not anticipated.

Next, the sample members were compared on the factor of age of clients with whom they worked. Of the psychiatrists, 25 worked with youth only, or with youth and adults; 27 worked with adults only. Of the clinical psychologists, 12 worked with youth only, or with youth and adults. Eleven worked with adults only. The two groups, those who worked with youth and those who worked with adults only, were compared on Items 1 to 6 as listed in Table I. These items were selected because they were mentioned the most often, 20 or more times, by the participants. The results of the comparison are presented in Table II.

The greatest differences within the ranks of the psychiatrists were on Item 4, interpersonal relationships, with eight of the 25 therapists (32 percent) who worked with youth and only four of the 27 (14.8 percent) who worked with adults suggesting it for curriculum inclusion; and on Item 6, sex and its normal manifestations, with ten of the 25 (40 percent) youth-oriented and five of the 27 (18.5 percent) adult-oriented therapists giving it emphasis.

The greatest differences within the ranks of the clinical psychologists were on Item 2, the emotions, with three of the 12 (25 percent) youth-oriented, in contrast to nine of the 11 (81 percent) adult-oriented therapists proposing it; and on Item 4, interpersonal relationships, with one of the 12 (8.3 percent) youth-oriented in contrast to seven of the 11 (63.0 percent) adult-oriented therapists suggesting it for curriculum inclusion.

When the psychiatrists and clinical psychologists were combined, only Item 6, sex and its normal manifestations, showed any sizeable difference in response. Thirteen of the 37 (35.1 percent) youth-oriented and seven of the 38 (18.4 percent) adult-oriented therapists suggested it for curriculum inclusion. The ratio is approximately two to one in favor of the youth-oriented group.

Curriculum-methods suggestions. When asked to identify the methods which should be employed to create a learning situation out of which the concepts of personality could "come alive", the participants leapt in many directions.

A number recommended the indirect approach. Fourteen hypothesized that with the teacher as the proper identification model, direct teaching would be unnecessary. All of these directed their recommendations at elementary, not secondary education. Ten believed that personality concepts should emanate from the "normal" curriculum per se.

Only six participants recommended that a course be added to the curriculum. The delineation of specific methods that were proposed follows:

TABLE I

CURRICULUM-CONTENT SUGGESTIONS

Psychological Concepts	Psychoanalysts (25)	Non-Psychoanalytical Psychiatrists (32)	Clinical Psychologists (25)	Total (82)	Median Age of Introduction
1. Defense Mechanisms compensation, pleasure-substitute mechanisms, introjection, repression, sublimation, conversion, projection	10	10	16	36	10. 8 years
2. The Emotions, Desires, Attitudes, Drives effect of emotion on logic, frustration tolerance, normalcy of desires but need to control, childhood fears and conflicts, need of balance in daily living, attitudes--how learned and changed, etc. , understanding the biological and social drives, etc.	9	7	14	30	10. 5 years
3. Symptomatic Reactions narcissism, reaction formation to rules regarding good and bad, ambivalence, rivalry, aggressivity, acting out as hostility, inferiority and over-striving, dependency	8	8	11	27	12. 3 years
4. Inter-personal Relationships relation of the individual to others, altruism--the aim of all therapies, ostracism and group identification, need to be accepted, status in the largest group that he can participate in, etc.	6	6	10	22	9. 0 years
5. Concept of Self and Self-Understanding need to understand self, need for self-esteem and acceptance, need to accept responsibility for ones own behavior, understanding of reality vs. pleasure principle, individuality vs. conformity, etc.	9	5	7	21	13. 7 years
6. Sex and Its Normal Manifestations how one becomes a man or woman, acceptance of the sexual nature of man, reproduction, parenthood, masculinity-feminity	10	5	5	20	13. 2 years

7. *Topographical Concepts* *the conscious a small part of our lives, the unconscious, the preconscious*	5	6	3	14	13. 1 years
8. Structural Concepts impulse or id, super ego, ego	4	4	4	12	13. 5 years
9. Mental Health and Illness, in General concept of adjustment, personality can be modified, maturity is never complete, various forms of psychotherapy, importance of learning in mental health, etc.	2	3	6	11	13. 0 years
10. Developmental Stages of Growth and Understanding maturation, Piaget's developmental stages, stages of development of personality functioning, etc.	4	--	4	8	12. 5 years
11. Sociological Concepts crime, psychology of the group, the community and emotional illness, mass psychology, cultural relativism, etc.	3	1	3	7	10. 5 years
12. Family Relationships sibling rivalry, family living, parent surrogates and symbolism, ambivalence toward parents, etc.	3	2	2	7	11. 5 years
13. Multiplicity of Causuality of Poor Mental Health complex underlying factors, causes of anxiety and frustration, felt need to be punished, psychic determinism, dynamics of conflicts, etc.	3	1	2	6	14. 5 years
14. Values and Ethical Considerations need for purpose in living, moral ethical considerations, etc.	2	1	2	5	12. 5 years
15. Abnormal Symptoms alcoholism as illness, character disorders, abnormal dealing with conflicts, sexual abnormalities	2	--	1	3	13. 5 years
Total	80	59	90	229	

TABLE II

RESPONSES COMPARED ON THE BASIS OF AGE OF CLIENTS SERVED

	All Psychiatrists		Clinical Psychologists		Totals	
	Youth Only or Youth and Adults (25)	Adults Only (27)	Youth Only or Youth and Adults (12)	Adults Only (11)	Youth Only or Youth and Adults (37)	Adults Only (38)
1. Defense Mechanisms	9	10	6	8	15	18
2. The Emotions	9	6	3	9	12	15
3. Symptomatic Reactions	8	8	4	6	12	14
4. Inter-Personal Relationships	8	4	1	7	9	11
5. Concept of Self, etc.	5	8	3	4	8	12
6. Sex and its Manifestations	10	5	3	2	13	7
Totals	49	41	20	36	69	77

Suggested Teaching Methods

1.	Teaching by examples and illustrations	61
	A. case studies B. autobiographies C. personal anecdotes D. classroom examples E. moral stories	
2.	Large group discussion	53
3.	The lecture	22
4.	Movies and film strips	20
5.	Cartoons	14
6.	Definitions of terms	10
7.	Role playing	7
8.	Small group sessions	6
9.	Dynamic interpretations	4
10.	Play therapy	2
11.	Puppets	2
12.	Others	15
	Total	216

As may be noted, the first choice of teaching method was by examples and illustrations. Next in line were the verbal methods of discussion and the lecture. Then follow such visual methods as movies, film strips, and cartoons. The suggestions were well thought out, occasionally novel, and more often than not in the direction of simulated life situations.

Teacher-education suggestions. In order to elicit responses that might be significant for those who work in teacher-education programs, the investigator, as indicated earlier, included the following item in the questionnaire:

> Assuming that a teacher of the afore-mentioned concepts is a reasonably well-adjusted, emotionally mature individual, what specialized training in the field of personality dynamics do you believe would be a minimum pre-requisite for his teaching role?

To say the least, the responses were many and varied. And they ran the gamut from no extra preparation to extensive additional preparation. A listing of the various suggestions and the number of respondents who made them follows:

Suggestions for Teacher Preparation

1. An M.D. is required with a specialty in psychiatry 1
2. A Ph.D. is required with a major in psychology, sociology, or relate fields 3
3. An M.S. is required with a major emphasis in psychology and related fields 9
4. The teacher must first undergo a personal analysis 2
5. The teacher must understand the dynamics of self and others 43
   Basic training, but method not stated 17
   Courses in dynamics 13
   Group therapy seminars 13
6. The teacher should have clinical experience with normal and institutionalized abnormal children, youth, and adults 24
7. He should enroll in courses in psychology of various types: general, child, abnormal, developmental 20
8. He should enroll in courses in sociology and anthropology 7
9. He should enroll in miscellaneous other courses 5
10. He should read Freud, Sullivan, Piaget, Adler and others 7
11. He should be personally selected by a psychiatrist or group of experts 3
12. He should be enrolled in an entirely new teacher-education curriculum with the emphasis on the understanding of self and others 7

Of significance in the recommendations is the obvious point of view that if teachers are to help others to improve their mental health, they, themselves, must be moderately well informed in the field of mental health. Yet within this frame of reference, moderation prevailed. Only 15 (note Items 1-4 above), made proposals which could be interpreted as requiring extensive extra work. The remainder seemed to think that with a different emphasis in the already existing teacher education curriculum or with a few additions to it, reasonably mature teachers could get ready to "teach about" selected dynamic concepts of personality.

Summary

The purpose of the research was to discover if specialists in the field of mental health believed that elementary and secondary school personnel could, with profit to children and youth, incorporate selected mental health concepts directly into the curriculum and approach them frontally. In other words, to what extent might the teaching about mental health concepts at the appropriate readiness level act as a preventive or alleviator of emotional disorders at a later time?

The sample consisted of 373 psychiatrists and 80 clinical psychologists selected at random from the respective professional registers. All had M. D. or Ph. D. Degrees, respectively. Of the 82 active respondents (one made only a token gesture), 57 were psychiatrists and 25 clinical psychologists. Of the 57 psychiatrists, 25 were psychoanalytically oriented. Thirty-two were not.

The 82 individuals had spent a mean 11. 6 years in therapy. The data-gathering device was an open-ended questionnaire. Regarding their attitudes toward the project, 48 were positive (58. 5 percent), ten were ambivalent (12. 2 percent), 23 were negative (28 percent), and one was too vague to classify.

Strongest curriculum support was afforded the following five areas of personality dynamics, in the order listed: the defense mechanisms; the emotions, desires, and attitudes; selected symptomatic reactions; inter-personal relationships; and the concept of self. In first priority with the psychoanalysts was sex in its normal manifestations; with the other psychiatrists and clinical psychologists, the defense mechanisms.

The factors of age level of clients and years of therapy made for some but not considerable differences in the responses.

In the area of teaching method, the sample members placed illustrations and examples first in priority, followed by group discussions, the lecture, movies, and cartoons.

In the area of teacher training, 15 of the 82 sample members believed that extensive work beyond the B. A. would be required if teachers were to teach selected dynamic concepts directly. By far the vast majority of those who responded believed that with minor additions to, or a shift of, emphasis within the existing teacher-education curricula, teachers could get ready for their proposed role in mental health.

Conclusion

The primary conclusion that can be reached from the study is that the majority of the limited sample of 82 experts in mental health support the idea that a frontal approach of some sort can be taken in the curriculum to selected dynamics of personality. More research of an extensive and intensive nature is needed, however, to support this conclusion and to refine the specific implications of it.

IT SAYS . . . . Continued. . . .

EXPERTS IN MENTAL HEALTH support the idea that a frontal approach of some sort can be taken in the school curriculum to mental health, Inlow says, but more research is needed to develop specific applications.

HOW DO PRINCIPALS BEHAVE? Cooper reports that in 62 percent of the total behaviors she studied, the principal was effective; but that in the area of personal characteristics and professional standards, the principal was ineffective in 65 percent of the behaviors observed. She suggests further research to determine whether such ineffectiveness is an inherent weakness in the preparation program.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 8, April 1963)

Changes in Self-Awareness During the High School Years: A Study Of Mental Health Using Paper-and-Pencil Tests

JOHN E. NICKOLS, Jr.*
Washington, D. C.**

MAN HAS ALWAYS been inescapably drawn to study his own significance in relation to other men. The early dictum of Protagoras that man is the measure of all things was a declaration of man's freedom from the inevitable law of the universe. Man was, thus, charged with mastering his own fate and satisfying his own desires. Subjective reality was illuminated as the objective reality of the perceiver; man became conscious of a connection between the individual's inner world and his own behavior (1). With this new awareness, the mind of man was destined to become a respectable topic for objective evaluation.

Socrates contributed to an effective understanding of the individual as a member of a group by propounding that man need only attain knowledge in order to choose the right thing (his belief in the individual's ability and integrity).(1) This gave birth to the corollary belief that accumulation of knowledge about oneself and others is imperative for correct action in relation to the group. It was prophetic that Socrates focused upon the individual as the means to knowledge about man and the group. His phrase, "Know thyself," has become axiomatic to modern specialists in human relations.

Today, it is commonly accepted that an intellectual knowledge of an individual's assets and liabilities influences his level of adjustment. Shaffer (2) refers to the "well-adjusted individual" as one who has objective knowledge of his own conduct and of external circumstances. The person who understands his own behavior can be viewed in direct contrast to the person dominated by self-deceit and rationalization. The aware person can acknowledge error and failure. He is in a position to select alternative modes of action and to learn more about himself and his relationships with others — he can profit from experience.

Sound self-appraisal allows us to be productive in relation to others. Without this situational awareness, it becomes only too easy to compound aberration; i.e., to construct an erroneous personal reality.

The purpose of this research was to investigate whether the adolescent's extreme oscillation of mood could confuse him to the point of losing an objective picture of his relative mental health status; or whether this variability is by nature experimental, leading to acquisition of more useful knowledge about the self. A secondary purpose is to investigate the utility of the California Mental Health Analysis as a paper-and-pencil type of evaluative criterion in research. This paper also contributes information about the utility of longitudinal versus cross-sectional designs.

The null hypothesis was that no significant change would occur in the adolescent's ability to judge his mental health standing accurately in relation to peers, during a three-year period from the freshman through the senior year in high school.

---

* Acknowledgements are made to Dr. Marvin Meyer, Cedar Rapids Public Schools, Iowa, for his kindness in administering the follow-up tests of the longitudinal study; to Marcia Nickols, Child Guidance Service, Lexington, Kentucky, for her editing suggestions; and to Dr. Marvin Strauss, Charleston Guidance Clinic, West Virginia, for his suggestions in clarifying several technical presentations.

** United States Public Health Service Outpatient Clinic, South Building.

Cross-Sectional Study *

Waraday and Nickols (3) observed sixty subjects of a parochial school, centrally located to serve children from all sections of a large city. Thirty freshmen were matched with thirty senior girls on the basis of Otis IQ, race, home situation, and length of school attendance. Thus, the research Ss consisted of Caucasians who had both English speaking parents living in the home, who had never failed or skipped a grade in school, and who were at least lower-middle class in socio-economic status. Mean IQ values were 102 for the freshmen and 103 for the seniors.

The evaluative criteria consisted of the California Mental Health Analysis (CMHA) and a specially devised personality questionnaire (PQ). The PQ consisted of ten items designed to elicit self-ratings relative to each of the ten subtest measures on the CMHA. The ten items were presented in simple words, and the Ss were asked to rate themselves in the highest to lowest fifths for girls in their classes. (Ratings made on items 2 and 5 were reversed; the lowest rating is equivalent to a rating designating the highest fifth on other items.) The PQ begins with comprehensive instructions: a sample item on body height is provided to orient the Ss to the rating procedure. The key terms of the ten items of the Waraday-Nickols Self-Image Scale are presented in Table I.

Procedures consisted of testing the total population of thirty-nine freshman and thirty senior girls in attendance on a randomly selected school day. The CMHA and PQ were administered in counterbalanced order. Precautions were taken to hold testing and motivation variables as constant as possible. Tests were administered during the last three weeks of the school year.

Five types of data were evaluated statistically. A self-rating (SR) value of 1 to 5 was obtained from each PQ item. For items one through nine, S received a high value of 5 for rating herself in the most self-enhancing fifth, and a low value of 1 for rating herself in the least self-enhancing fifth of her class. For item ten, an average of each S's two ratings was used. A total score for self-ratings (SR_t) was determined for each S. This was done by summing the ten SR values obtained from her PQ rating sheet. Out of a possible score of 50.0, SR_t means were 36.9 for the freshmen and 37.4 for the seniors, the difference not being significant. The differences between the groups did not deviate from a chance distribution for the ten PQ items evaluated individually.

A class standing (CS) value of 1 to 5 was devised to represent a CMHA subtest score relative to each S's classmates. The method of cumulative frequencies to the mid-point was used to convert the raw scores for each subtest into percentile ranks. For each subtest, Ss with a class percentile of 80 to 99 received a high CS rating of 5, and a percentile of 0 to 19 received a low rating of 1. An overall class standing (CS_t) value was determined for each S by summing her ten (subtest) CS values into a total score. The average CS_t for each group was approximately 10 (subtests) × 3, or 30.0.

A normative standing (NS) value of 1 to 5 was assigned to each S on the basis of standardization norms reported in the test manual. For each CMHA subtest, a percentile norm of 80 to 99 received a high rating of 5; and 0 to 19 a low rating of 1. A normative standing total (NS_t) value was determined for each S by summing the ten NS values to represent an overall CMHA standing. This averaged 35.8 for the seniors and 31.7 for the freshmen, the difference being statistically significant at the .05 level of confidence. For both groups, these NS means were higher than the expected (CS_t) means of 30.0; the difference was statistically significant for the seniors.

NS's were included for three reasons. The small sample might be subject to deviations from a normal distribution; it was not possible to predict the extent a S's SR might be influenced by unintentional incorporation of comparisons with peers attending different schools; and the use of class percentiles, alone, obscures the numerical contribution of any one raw score.

The seniors achieved a mean of 151.5 on total CMHA raw scores (or $CMHA_t$); this was significantly greater than the mean of 138.7 achieved by the freshmen. The mean NS_t score of 35.8 for the seniors was significantly closer to their SR_t mean of 37.4; relative to the respective means of 31.7 and 36.9 achieved by the freshmen. This significant difference does not necessarily demonstrate that the individual senior possesses greater insight into her relative mental health. The difference represents an overall average, which obscures individual performances. The mean differences under discussion only support this prediction: while freshmen and seniors tend to rate themselves at similar mental health levels, the average mental health scores achieved by a group of seniors should correspond more closely with the average level at which they tend to rate themselves.

A class error (CE) value of 0 to 4 was determined to correspond to each of the ten subtests. This CE value represents the extent to which the SR value differs from a corresponding CS value; i.e., disregarding the direction of a difference, either the SR value was subtracted from its corresponding CS

---

* Conducted by Sherman Waraday, Chicago, Illinois, and the author, who are deeply indebted to the officials and students of a parochial school for their cooperation in making this project feasible; and to Dr. Harry Moore, University of Denver, for his services as advisor for this study.

TABLE I

SHOWING KEY PHRASINGS USED IN THE WARADY-NICKOLS SELF-IMAGE SCALE, CONSTRUCTED TO CORRESPOND TO THE TEN SUBTESTS OF THE CALIFORNIA MENTAL HEALTH ANALYSIS

CMHA Sub-tests*	Key Phrasings of Self-Rating Items**
Liabilities	
A. Behavioral Immaturity	1. How grown-up you are, mentally.
B. Emotional Instability	2. How sensitive you are. How much things and situations bother or disturb you.
C. Feelings of Inadequacy	3. How well you think you can do things like other girls in your class.
D. Physical Defects	4. Advantages you have because of facial and body appearance.
E. Nervous Manifestations	5. How nervous others see you (and how nervous you think you are).
Assets	
A. Close Personal Relationships	6. How well-off you are because of friends in whom you can confide, who respect you, and who make warm friendships.
B. Inter-Personal Skills	7. How well you overcome selfishness to make friends, to make others happy, and to take part in things with them.
C. Social Participation	8. How willing and successful you are in helping with and enjoying group activities.
D. Satisfying Work and Recreation	9. How good a feeling work and recreation give you.
E. Adequate Outlook and Goals	10a. How much others approve of your ambitions. b. How satisfied you are with your progress toward these ambitions.

* CMHA sub-tests were used to obtain scores for determining class- and normative-standings, or rankings.

** Questionnaire items were used to determine self-ratings.

value, or the CS value was subtracted from its corresponding SR value. A class error total (CE_t) was obtained merely by summing all ten differences (derived from each of the ten CMHA subtests and ten PQ items) for each S. A perfect CE_t is 0.0. CE_t means were 12.4 for the freshmen and 12.9 for the seniors, the difference being negligible statistically. None of the ten subtests revealed statistically significant differences between the groups.

A normative error (NE) also ranged from 0 through 4 on each subtest and for each S. This is the difference between each NS value and its respective SR value. A normative error total (NE_t) was obtained merely by summing the ten NE values for each individual. NE_t means were 11.8 for the freshmen and 10.5 for the seniors, with no significant differences between the groups ($t_d = 1.6$).

Several coefficients of correlation were calculated in relation to the above results. Significant coefficients of .55 for the freshmen and .63 for the seniors were obtained between CS_t and SR_t; the difference is not significant. Significant coefficients of .86 for the freshmen and .60 for the seniors were obtained between SR_t and NS_t. A difference-t of 2.02 is just barely short of the 2.04 needed for significance. (This high t_d can also be rejected on the basis that the corresponding difference is not significant in the subsequently observed longitudinal study.) Significant coefficients of -.71 for the freshmen and -.68 for the seniors were obtained between CS_t and CE_t. This difference is not significant; and coefficients for the ten subtests did not deviate significantly from a chance distribution.

Otis (self-administered) IQ's were correlated with $CMHA_t$, CE_t, and SR_t values. These correlations were all mildly favorable. No significant or consistent differences were observed between the groups on these variables.

In conclusion, the seniors as a group demonstrated significantly greater mental health scores than was the case for the freshmen. Students of better mental health seemed to possess greater insight into their personality assets and liabilities, as measured in this study. However, this situation did not seem to undergo significant change between the freshman and senior years of high school.

The results failed to support significant differences between the groups for self-ratings (SR_t values) and errors of judgment (NE_t values). This, and the fact that the seniors achieved significantly higher mental health scores, suggests that a comparatively greater number of seniors could underestimate themselves in relation to their peers. Perhaps this is associated with achievement of a greater sense of having been accepted by peers. In this respect, a greater number of freshmen might have been striving to look like "other kids" (as if playing a role to become seniors); whereas a greater number of seniors could somewhat relax their drive to "appear acceptable."

Correlation results suggested that a significant number of freshman and senior girls tended to function as if possessing realistic assumptions about their relative status in mental health assets and liabilities. At best, IQ seemed only mildly related to mental health scores, self-ratings, and accuracy in rating the self on mental health assets and liabilities.

Longitudinal Study

Three years after the cross-sectional study, twenty-two of the thirty-nine freshman subjects tested originally were available for re-evaluation with the CMHA and PQ. For these white, female Ss, the range in (self-administered) Otis IQ was 87 to 113. The mean was 102.2.

Although a different examiner was used for the follow-up testing, the procedures consisted of an attempt to reestablish the original testing conditions as nearly as possible. Test-retest coefficients of reliability were .59 for the total CMHA raw scores ($CMHA_t$); .52 for the sum PQ ratings (SR_t); .49 for total class error (CE_t); and .43 for the total normative error (NE_t). The means for $CMHA_t$ were 146.8 for the freshmen, the percentile being approximately 43; and 161.8 for the seniors, the percentile being approximately 70. This difference is highly significant ($t_d = 4.6$). SR_t means were 35.9 for the freshmen and 38.8 for the seniors. This difference is also highly significant ($t_d = 3.1$). SR and NS, then, increased significantly. CS, of course, could not change significantly.

These results suggest both that the seniors tended to rate themselves higher and that their mental health scores evidenced a corresponding increase.

CE did not change significantly, while the mean NE_t decreased 0.8, significant from the initial freshmen mean of 11.4 ($t_d = 2.6$). Even the mean group error ($SR_t - NS_t$, for each individual), decreased significantly from 7.6 to 3.7 ($t_d = 3.1$). These results suggest that, if standardization norms are used, the seniors were significantly better able to rate themselves in accordance with their relative mental health on the criteria used. This improvement is significant for the seniors, as individuals and as a group.

Coefficients of correlation were determined between SR and CS, and between SR and NS. For both CS and NS, nine correlations for the ten CMHA subtests (assets and liabilities) were positive for each group. In each comparison, the nine correlations were higher for the seniors, but seven of these differences were not significant for either CS or NS comparisons. Coefficients for the seniors were significantly higher on "feelings of inadequacy," between SR and CS ($t_d = 2.2$); and on "behavioral immaturity (freedom from)," between SR and NS ($t_d = 3.8$). "Close personal relationships" was the only subtest consistently significant for both CS and

NS. The seniors dropped significantly on this subtest ($t_d = 5.0$ between SR and CS, and 2.2 between SR and NS).

The ten subtest coefficients between SR and CS averaged .36 for the freshmen and .44 for the seniors. For SR and NS, the averages were .30 for the freshmen and .46 for the seniors. Although the seniors were higher on both averaged r's, neither difference is significant. Only the fact that the seniors were higher on nine of the ten subtests is significant. Coefficients between SR_t and CS_t were .70 for the freshmen and .76 for the seniors. Coefficients between SR_t and NS_t were .69 for the freshmen and .83 for the seniors. Both of these comparisons show insignificantly higher correlations for the seniors.

These results, and those for NE_t's, suggest a tendency for the individual senior to rate herself closer to her tested mental health standing. However, the seniors as a group might vary markedly in over- and underestimating themselves. The subtest correlations suggest that the seniors are significantly less able to rank their relative standings on "close personal relationships" (possibly they are also less aware of their achievement in having formed such relationships); however, they may be better able to rate themselves on "behavioral immaturity" and "feelings of inadequacy."

Exploratory coefficients of correlation were also determined between several types of initial scores and changes in SR_t and CMHA_t scores. No significant correlations were observed between SR_t gains and initial values for CMHA_t, SR_t, NE_t, or (Otis) IQ. Significant correlations were observed between CMHA_t gains and initial values for CMHA_t ($r = -.42$), NE_t ($r = .47$), and IQ ($r = .39$). The coefficient was .00 between CMHA_t gains and initial SR_t values.

These results indicate that gains in CMHA raw scores tended to be somewhat greater for students of higher intelligence and/or of initially lower abilities to "predict" their tested mental health standings. Students of higher mental health as freshmen manifested a slight trend for the least gain in mental health test scores; the inherent ceiling limitations of the evaluative criteria might be significant here. Changes in self-ratings did not seem significantly related to initial mental health scores, self-ratings, accuracy scores, or initial IQ.

Discussion

To evaluate the results of these studies realistically, it is also important to evaluate the measurement criteria. The personality questionnaire devised for this study had a reliability coefficient of .52. On the one hand, this speaks well for the self-ratings, in view of the fact that total raw scores on the California Mental Health Analysis (CMHA) were associated with a reliability coefficient of .59. The reliability coefficients suggest that sums for the ten self-rating items (or for the eleven questions) of the personality questionnaire yield stability which is comparable to that achieved with the 200 items of the CMHA. The fact that the reliability coefficient is .49 for total class error, and that the coefficient for the sum normative error is .43, further suggest that the stability of the differences between self-ratings and CMHA standings does not differ markedly from the stability of either measurement criterion. Thus, it seems as permissible to use self-ratings, class error, or normative error as to use CMHA raw scores as the evaluative criteria in this exploration into the development of self-conception.

On the other hand, test-retest coefficients for all four sets of measurement are relatively low. The obtained reliability coefficient for total CMHA raw score of .59 is substantially below that of .93 calculated by the Richardson-Kuder method and reported in the test manual. (The standard deviations for the longitudinal study were 16.1 for the freshmen and 16.7 for the seniors. The standard deviation was 19.0 for the thirty freshmen used in the cross-sectional study. These values should be close enough to the standard deviation of 21.5 reported in the test manual to suggest a fair comparison.

Since, for this investigation, validity of evaluating changes in self-conception is dependent upon the accuracy of measuring mental health, the low CMHA test-retest reliability permits one to doubt the conclusiveness of the evaluative criteria. Subsequent discussion is presented with these qualifications in mind.

The most outstanding finding of this investigation is that the seniors achieved significantly higher scores on the CMHA than the freshmen. The differences were approximately 18 percentile points for the cross-sectional and 27 percentile points for the longitudinal study. The latter study further suggests that intelligence is positively related, and initial total scores on CMHA negatively related, to CMHA total score gains made during the three year period. (Because of the use of independent Ss, these comparisons were not possible in the cross-sectional study.)

These results need not suggest that only freshmen of higher intelligence and initially lower mental health tend to gain the most mental health scores. It might be that those Ss with higher mental health as freshmen have the least potential for improvement on the CMHA as seniors. Further, the CMHA might have inherent ceiling limitations; i.e., the test might not discriminate well at the upper range of mental health, and the true-and-false type of item might allow a high degree of error to intrude into the test scores.

It is interesting that initial ability to rate oneself accurately on mental health characteristics was negatively related to CMHA score gains. This might suggest that the freshmen least aware of themselves have the greatest potential for improvement during the last three years of high school.

In general, then, the last three years of high school might be accompanied by circumstances which affect the child with lower mental health and/or with less cognitive insight into his assets and liabilities, as measured in this study. Conceivably, these conditions could represent a problem-solving situation for the "less healthy" and/or "less insightful" child; i.e., he might encounter greater circumstantial pressures to resolve his problems in living. The student of higher intelligence might be in the better position to respond favorably.

Self-ratings increased a statistically significant 2.9 points for the longitudinal study, whereas the difference between the groups was negligible for the cross-sectional study. The freshmen and seniors for both studies tended to rate themselves at an average of about five to nine points above the expected mean of 30.0 for the sum of each person's ten self-ratings; this might suggest that the students as a group shied away from thinking of themselves as below average in mental health. The fact that only the longitudinal study revealed significant increases in self-rating values might be associated with the superiority of the longitudinal approach; e.g., this approach eliminates the need to match samples and allows the reliability of the evaluative criterion to be taken into consideration in the statistical treatments of the results. In the senior year, overestimation of oneself seems less unrealistic, in view of the fact that the seniors as a group tended to fall above the standardization median in mental health scores.

The correlations between self-ratings and rankings on CMHA scores, both in terms of class and normative comparisons, were highly significant for the freshmen and seniors of both studies. Although the coefficients were generally higher for the seniors, the differences between the groups failed to reach significance. Coefficients for the ten assets and liabilities (derived from the CMHA subtests), also tended to favor the seniors, but the difference-t's were generally insignificant. The one clear-cut exception was that the seniors consistently fell significantly below the freshmen on "close personal relationships." Probable exceptions were that the seniors gained significantly in "behavioral immaturity" and "feelings of inadequacy," either in terms of class or standardization norms.

These results suggest that both freshmen and seniors tended to demonstrate somewhat realistic self-conceptions, while neither group tended to be significantly superior in the consistency with which self-ratings correspond with their obtained mental health scores. The three mental health factors suggesting a differential in self-conception were difficult to interpret because of lack of clarity in defining their relevance. If the differences are not related to errors in measuring mental health with these CMHA subtests, it might be a hint that the seniors tend to be better aware of how "grown-up" they are mentally, and how well they can "do things" in relation to their peers. However, the seniors seem less aware of their relative assets in "close personal relationships"; i.e., in having achieved warm frienships, respect, and trusted companions from among their peers. Perhaps the seniors are developing more social contacts outside of their peer group. If so, they would be trying to "find themselves" in relation to their new associates.

Correlations failed to reveal significant relationships between changes in self-ratings and initial values on the IQ, CMHA, self-conception (error in judgment), and self-rating variables. This suggests that change in one's opinion of himself is not generally related to intelligence, previous mental health scores, previous errors in judging mental health characteristics, or previous opinions about one's mental health assets and liabilities.

Class error (discrepancies between self-ratings and tested mental health standings at the level of class comparisons), did not tend to change significantly. Highly significant negative correlations between mental health standing and class error suggest that mental health is positively related to self-awareness, but this relationship did not change significantly for the last three years of high school.

Normative error in self-conception (discrepancies determined by the use of standardization norms), tended to decrease during the last three years of high school. Allowing for the fact that this difference was only significant for the longitudinal study, these results might indicate that the seniors as a group tend to operate with fewer erroneous assumptions about their general mental health.

In conclusion, the null hypothesis cannot be rejected with sufficient confidence for the two major (class and normative) areas of comparison. The more adequate experimental designs might reveal greater overall accuracy among the seniors in estimating mental health assets and liabilities, relative to their respective peer group. It might also be found that, while the seniors over- and underrate themselves almost as frequently as the freshmen, the seniors will more closely approach their objectively determined average ranking. This hypothesis seems to apply only if standardization, rather than class, norms are used to determine relative mental health standings.

Summary

The primary purpose of this project was to explore the effects of the high school years upon mental health scores and related self-awareness. The cross-sectional study utilized matched groups of thirty freshman and thirty senior girls. For the longitudinal study, twenty-two of the freshmen tested originally were followed-up three years later.

The evaluative criterion consisted of two basic scales. The California Mental Health Analysis (CMHA) was used to obtain objective mental health scores. A rating scale was devised to obtain self-appraisals on ten variables which might correspond to the assets and liabilities of the California Scale.

The reliability coefficient of .59 for total scores obtained with the California Mental Health Analysis over a three-year period was disappointing, particularly in view of the fact that the personality questionnaire of only eleven items had a coefficient of .52. The fact that an average correlation of over .70 was found between total CMHA scores and summations for the self-ratings suggests that scores on the California Scale are strongly influenced by subjective attitudes about the self.

Although some of the results cast doubts upon the usefulness of paper-and-pencil tests in exacting research, certain speculative hypotheses might be extracted for future study. The results of both studies clearly suggest that the seniors can be expected to achieve significantly higher mental health scores. Descriptive variables positively related to mental health scores seem to include (a) the level at which the individual tends to rate himself, and (b) his accuracy in rating his relative mental health status. Freshman intelligence might be positively related to the degree of improvement made in mental health scores between the freshman and senior years. Initial mental health scores and initial ability to assess oneself accurately correlated negatively with gains in mental health scores for the subjects used in the longitudinal study.

Only the longitudinal study suggests that under some conditions the seniors might be more aware of their relative rank in mental health asset and liability scores. While both the freshmen and seniors tended to over-rate themselves, and while neither group seemed able to avoid marked variance in over- and under-estimating themselves on the ten assets and liabilities studied, to a significant extent both groups seemed to possess somewhat realistic assumptions about their overall mental health, as measured in this study. Students with better mental health scores seemed to be somewhat more accurate in appraising themselves.

REFERENCES

1. Frost, S. E., Jr. Basic Teachings of the Great Philosophers (New York: Barnes and Noble, Inc., 1947).
2. Shaffer, L. F. The Psychology of Adjustment (New York: Houghton Mifflin Co., 1936), pp. 536-7.
3. Warady, Sherman and Nickols, J. E., Jr. An Exploratory Investigation into the Development of Self-Conception during the High School Years, unpublished master's thesis, University of Denver, 1952.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 8, April 1963)

An Analysis of the Quality of the Behaviors of Principals as Observed and Reported in Six Critical Incident Studies

BERNICE COOPER
University of Georgia

THE CRITICAL requirements for an effective principal have been identified through six studies (1, 2, 3, 5, 6, 7) conducted at the University of Georgia. It was the purpose of the present investigation to analyze the quality of the behaviors of the principal as observed and reported for the development of the critical requirements. The objectives of the analysis were as follows: 1) to determine whether the principal was more effective than expected by chance in certain behaviors, 2) to determine whether the principal was more ineffective than expected by chance in certain behaviors, and 3) to identify specific behaviors wherein the principal was ineffective.

The Six Studies

Procedure. The six studies, with the primary purpose of determining the critical requirements for the principal, were conducted by doctoral candidates of the University of Georgia through a grant-in-aid from the W. K. Kellogg Foundation from 1954-56. All of these studies employed the critical incident technique, a method of data-gathering which uses observed incidents of human behavior that make the difference in effective or ineffective performance. The respondent groups included principals, teachers, pupils, instructional supervisors, visiting teachers, school board members, and college instructors in teacher training institutions who have opportunities to observe the principal.

Frequency of effective and ineffective behaviors. A total of 8040 critical incidents, yielding 8649 critical behaviors, were reported by the seven publics. So far as the primary purpose of the six studies was concerned, the development of the list of critical requirements did not distinguish between effective and ineffective behaviors. In the categorization process, all behaviors became positive statements of critical requirements for effective administration. However, the reported incidents, according to the research design, were judged as effective or ineffective by the respondent. From Table I, it may be observed that, of the 8649 behaviors, 5338, or 61.72 percent were judged to be effective behaviors and 3311, or 38.28 percent were judged to be ineffective. Thus, in slightly more than three-fifths of the total behaviors which were deemed critical in the performance of the job, the principal was reported to be effective.

An Area of Ineffective Behaviors. The six investigators categorized the critical behaviors into general areas, ranging from five to eight areas per study.

From Table II, it may be noted that, in the majority of instances, the relationship between the areas and quality was significantly effective. However, all six studies reported significantly more ineffective than effective behaviors in the area of personal characteristics and professional standards than expected by chance. The import of this finding becomes even greater when it is noted that this particular area was the only significantly ineffective area according to four of the six studies. The remaining two each found a relationship between ineffectiveness and one other, but different, area.

From the combined studies, the areas encompassing personal characteristics and professional standards contained 978 behaviors, or approximately eleven percent of the total number of behaviors observed and reported. When the remaining 7671 behaviors were examined, the percentage of effective behaviors was considerably higher than the percentage of ineffective behaviors. Table III shows that 65.01 percent of the behaviors were effective as compared with 34.99 percent ineffective— or about 2 of 3 behaviors effective. The chi-square test was used to test the null hypothesis that effectiveness was unrelated to the areas of behaviors. The null hypothesis was discarded when P was found to be less than .01. It was concluded, then, that significantly more effective behaviors than expected by chance were reported in all the areas, less the area of personal characteristics and professional standards.

TABLE I

THE FREQUENCY OF INCIDENTS AND OF EFFECTIVE AND INEFFECTIVE BEHAVIORS FROM THE SIX STUDIES

Investigator	Number of Incidents	Effective Behaviors		Ineffective Behaviors		Total Behaviors	
		Number	Percent	Number	Percent	Number	Percent
1	1042	617	58.76	433	41.24	1050	100.
2	695	502	65.71	262	34.29	764	100.
3	1986	1531	77.09	455	22.91	1986	100.
4	2557	1533	52.02	1414	47.98	2947	100.
5	729	539	61.88	332	38.12	871	100.
6	1031	616	59.75	415	40.25	1031	100.
TOTAL	8040	5338	61.72	3311	38.28	8649	100.

TABLE II

THE FREQUENCY OF SIGNIFICANCE OF THE RELATIONSHIP BETWEEN AREAS AND QUALITY IN THE SIX ORIGINAL STUDIES

Investigator	Total Number of Areas in Study	Number of Areas		
		Significantly Effective	Significantly Ineffective	No Significant Difference
1	5	3	1	1
2	6	2	1	3
3	8	4	2	2
4	7	3	2	2
5	5	3	1	1
6	7	6	1	0

TABLE III

THE FREQUENCY OF EFFECTIVE AND INEFFECTIVE BEHAVIORS LESS THOSE IN THE AREA OF PERSONAL CHARACTERISTICS AND PROFESSIONAL STANDARDS

Investigator	Effective Behaviors		Ineffective Behaviors		Total Behaviors	
	Number	Percent	Number	Percent	Number	Percent
1	579	62.87	342	37.13	921	100.
2	450	69.12	201	30.88	651	100.
3	1381	79.46	357	20.54	1738	100.
4	1492	54.77	1232	45.23	2724	100.
5	506	68.56	232	31.44	738	100.
6	579	64.40	320	35.60	899	100.
TOTAL	4987	65.01	2684	34.99	7671	100.

TABLE IV

THE FREQUENCY OF EFFECTIVE AND INEFFECTIVE BEHAVIORS IN THE AREA OF PERSONAL CHARACTERISTICS AND PROFESSIONAL STANDARDS

Investigator	Effective Behaviors		Ineffective Behaviors		Total Behaviors	
	Number	Percent	Number	Percent	Number	Percent
1	38	29.46	91	70.54	129	100.
2	52	46.02	61	53.98	113	100.
3*	150	60.48	98	39.52	248	100.
4	41	18.39	182	81.61	223	100.
5	33	24.81	100	75.19	133	100.
6	37	28.03	95	71.97	132	100.
TOTAL	351	35.89	627	64.11	978	100.

*Expectancy was computed in terms of proportion of successful and unsucessful incidents reported in total number of incidents reported.

The reverse was found when only the behaviors in the area of personal characteristics and professional standards were considered. From Table IV it will be noted that, of 978 behaviors in this area, 627 behaviors, or 64.11 percent, were ineffective and 351, or 35.89 percent, were effective. Then, in approximately two of three behaviors related to this area, the principal was ineffective. The chi-square test, with P less than .01, showed that significantly more ineffective behaviors than expected by chance were reported in this area.

These findings indicated that further study of requirements in the area of personal characteristics and professional standards needed to be made in order to identify specifically ineffective behaviors.

An Analysis of Behaviors

The analysis of behaviors in the area of personal characteristics and professional standards included 1) the development of a composite list of behaviors in the area being analyzed, 2) a statistical analysis to determine the significance of the relationship between quality and the requirement, and 3) the identification of the requirements wherein the principal's behavior was more ineffective than expected by chance.

Composite list of requirements. From the six studies, a composite list of the requirements in the area of personal characteristics and professional standards was derived inductively. The studies listed 14, 18, 31, 21, 11, 18, respectively, or a total of 113 requirements in this area. For compilation purposes, the first sample of 25 requirements included a proportionate number from each study. The first sample of 25 requirements was examined; and identical requirements, or those so similar that for practical purposes they were the same, were grouped. This process was repeated with each new sample's being added to the groupings or new groupings formed. The behaviors encompassed in each requirement were checked constantly to insure proper interpretation of the original requirement. When all the samples had been processed, the statements were rewritten to be all-inclusive of the requirements pertaining to a common element. Thirty-three critical requirements in the area of personal characteristics and professional standards were thus established.

Then, of these 33 requirements, which particular ones, if any, were significantly more effective than expected by chance?

Statistical analysis. Statistical analyses were made of the data to determine whether a significant relationship existed between the requirement and the effectiveness or ineffectiveness of the behavior. The null hypothesis that the effectiveness or ineffectiveness of the behavior was unrelated to the requirement was tested by the chi-square test. To judge whether the obtained chi-square was significant, when P was greater than .05, chi-square was considered not significant, and the null hypothesis was retained. Conversely a P of .05 or less was interpreted to be "significant at the .05 level," and the null hypothesis was discarded; i.e., the hypothesis of independence was rejected.

Relationship between quality and requirements. Table V lists the requirements with the frequency of effective, ineffective and total behaviors and summarizes the results of the chi-square tests which determined the existence of a relationship between quality and the requirement. It will be noted that in seventeen tests, the null hypothesis was discarded. In four tests the principal was observed performing more effective behaviors than expected by chance. And, in 13 tests, the principal was observed performing more ineffective behaviors than expected by chance.

The following 13 requirements, then, become identified as the ones wherein the principal is most ineffective, according to these six studies:

1. Observes and encourages personal habits of good grooming, cleanliness, and orderliness.
2. Adheres to and upholds high moral and ethical standards in personal and professional life.
3. Displays and upholds social behavior compatible with accepted standards of conduct.
4. Does not let family or personal feelings influence school actions or decisions.
5. Is accurate and honest in personal and professional record-keeping and financial transactions.
6. Expresses his intentions openly and sincerely but refrains from threats or displays of authority.
7. Shows consideration for others and respects personal rights.
8. Criticizes, reprimands, or punishes the offender privately and justly.
9. Displays traits of punctuality, dependability, and truthfulness.
10. Uses clean speech in presence of students.
11. Apportions his time wisely and equitably among the various phases of the school program.
12. Demonstrates effectiveness in basic skills of oral and written communication.
13. Determines and fulfills the responsibilities which have been delegated to him, not letting other jobs interfere.

Conclusions and Recommendations

From the analysis, the following conclusions were made:

1. From the six original studies, in 61.72 percent of the total behaviors observed and reported,

TABLE V

THE RELATIONSHIP BETWEEN EFFECTIVENESS OR INEFFECTIVENESS AND THE REQUIREMENTS

	Behaviors					
Requirements	No Eff.	No In-eff.	Total	df	χ^2	P
1. * Observes and encourages personal habits of good grooming, cleanliness, and orderliness	1	15	16	1	10.56	P > .01
2. * Adheres to and upholds high moral and ethical standards in personal and professional life	9	49	58	1	26.224	P > .01
3. * Displays and upholds social behavior compatible with accepted standards of conduct	2	29	31	1	21.806	P > .01
4. Maintains emotional stability and evidences the ability to relax under pressure	9	8	17	1	—	—
5. Lends support to and participates actively in community and civic projects, church and youth activities, educational functions and organizations, but maintains balance between personal and professional responsibilities	29	7	36	1	12.250	P > .01
6. * Does not let family or personal feelings influence school actions or decisions	3	28	31	1	18.580	P > .01
7. * Is accurate and honest in personal and professional record-keeping and financial transactions	4	32	36	1	20.250	P > .01
8. Admits and assumes full responsibility for his mistakes and for actions of staff and students at school activities	7	36	43	1	1.824	.30 > P > .20
9. Faces problems courageously and acts promptly in calm manner	11	3	14	1	3.50	.10 > P > .05
10. Is patient, consistent and impartial when working with others, singly or in groups	31	28	59	1	.006	.95 > P > .90
11. * Expresses his intentions openly and sincerely but refrains from threats or displays of authority	0	15	15	1	13.066	P > .01
12. * Shows consideration for others and respects personal rights	3	12	15	1	4.266	.05 > P > .02
13. Evidences adaptability in making adjustments to situations and individuals or groups	7	12	19	1	.842	.50 > P > .30
14. Controls his temper	15	27	42	1	2.88	.10 > P > .05
15. Is generous with time, energy, and money to help others	35	10	45	1	12.8	P > .01
16. Exhibits a courteous and tactful manner	29	22	51	1	.706	.50 > P > .30
17. Gives credit, praise, or recognition to those responsible for jobs well done	13	15	28	1	.036	.90 > P > .80
18. Secures all facts in order to make open-minded and impartial decisions	7	6	13	1	—	—
19. * Criticizes, reprimands, or punishes the offender privately and justly	4	44	48	1	31.688	P > .01
20. * Displays traits of punctuality, dependability, and truthfulness	17	73	90	1	33.612	P > .01
21. * Uses clean speech in presence of students	0	10	10	1	8.10	P > .01
22. Obtains personal information in objective manner and treats it confidentially	3	5	8	1	.125	.80 > P > .70
23. Employs diplomacy and wise judgment in dealing with requests and demands of parents	4	10	14	1	1.786	.20 > P > .10
24. Makes advance plans which are presented through an effective communication system	7	12	19	1	.842	.50 > P > .30
25. * Apportions his time wisely and equitably among the various phases of the school program	6	17	23	1	4.348	.05 > P > .02
26. Adequately plans for and presides skillfully at meetings	8	14	22	1	1.136	.30 > P > .20
27. * Demonstrates effectiveness in basic skills of oral and written communication	6	17	23	1	4.348	.05 > P > .02
28. Demonstrates willingness to perform beyond the call of duty	7	7	14	1	.072	.80 > P > .70
29. Demonstrates willingness to do the same things that he may ask others to do	14	19	33	1	.484	.50 > P > .30
30. Shows a willingness to assume extra duties in exigent circumstances	21	1	22	1	16.410	P > .01
31. * Determines and fulfills the responsibilities which have been delegated to him, not letting other jobs interfere	12	36	48	1	11.02	P > .01
32. Maintains high standards and supports the policies which he believes are good for the school and community	21	7	28	1	6.036	.02 > P > .01
33. Is an effective teacher	6	1	7	1	2.286	.20 > P > .10
Total	351	627	978			

*A significant relationship between ineffectiveness and the requirement.

the principal was effective.

2. In the area of personal characteristics and professional standards, the principal was reported to be ineffective in 65.01 percent of the behaviors observed and reported.

3. The relationship between effectiveness and behaviors in all areas other than the area of personal characteristics and professional standards was significant at the .01 level.

4. The relationship between ineffectiveness and behaviors in the area of personal characteristics and professional standards was significant at the .01 level.

5. Of 33 requirements in the area of personal characteristics and professional standards, 13 were found to have a significant relationship between ineffectiveness and the requirement.

In terms of the conclusions, the following recommendations are made:

1. Further study of the requirements in the area of personal characteristics and professional standards should be made to determine whether the ineffectiveness is an inherent weakness of the preparation program.

2. Attention should be given to the findings of this study in the screening process and in the preservice and in-service preparation programs for principals.

REFERENCES

1. Bice, Herbert. Critical Requirements for the Principalship as Observed and Reported by School Board Members in Georgia, unpublished Doctor's Dissertation, University of Georgia (1955), 143 pp.
2. Cooper, Bernice. The Critical Requirements for the Principalship Based upon an Analysis of Critical Incidents Reported by Instructional Supervisors and Visiting Teachers, unpublished Doctor's Dissertation, University of Georgia (1956), 200 pp.
3. Elliott, Frank Samuel. Requirements for Public School Principalship in Georgia as Determined from Student Observations of the Actions of Principals, unpublished Doctor's Dissertation, University of Georgia (1955), 158 pp.
4. Flanagan, John C. "The Critical Incident Technique," Psychological Bulletin, LI (July, 1954), pp. 327-58.
5. Phillips, Herbert E. Critical Requirements for the Principalship in Georgia as Observed by Teachers, unpublished Doctor's Dissertation, University of Georgia (1954) 160 pp.
6. Robins, Gerald Burns. Critical Requirements for the Principalship: A Study of Observed Behaviors of Principals in Georgia, unpublished Doctor's Dissertation, University of Georgia (1954), 140 pp.
7. Tyson, Ralph K. The Critical Requirements for the Principalship from an Analysis of Critical Incidents as Reported by Observers in Teacher Education Institutions, unpublished Doctor's Dissertation, University of Georgia (1955), 166 pp.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 8, April 1963)

Academic Achievement and Adjustment of Underage and Overage Third Graders

MARIAN L. CARROLL
State University College
Buffalo, New York

INTEREST IN this study resulted from hearing primary grade teachers frequently explain a child's low academic achievement with an observation such as, "He is so young," or "She is too immature for the work." A few recent investigations have suggested that mental age and intelligence quotient can be over-emphasized as determiners of the child's early school success. Certain other factors, both physical and psychological, related to the child's developmental status, appear to deter or to facilitate the use he can make of his learning opportunities.

Among studies which have produced evidence of this kind was a comparison of the achievement of 73 pairs of Kansas City public school children, matched on intelligence quotient, sex and, in about two-thirds of the cases, the school first attended. One child of each pair had entered the fall term of kindergarten following a fifth birthday in January or February, while the child with whom he was paired did not reach the age of five until November or December after enrollment in school. Data from cumulative records showed that the overage group, from kindergarten through grade ten, had made significantly higher marks in school subjects, significantly higher achievement test scores in reading, arithmetic and social studies, received significantly higher ratings from teachers on personal traits, and were significantly more successful in moving regularly from grade to grade.[1]* A study reported by Carter compared the achievement of 100 children who had entered first grade in the Austin Public Schools in September 1947.[2] Half of these children were under and half over six when they enrolled in grade one. Matching was done on the basis of sex and intelligence quotient. Achievement test scores carefully recorded throughout elementary school provided data for comparison. Findings suggested that the chronologically older child had an advantage over the younger one throughout elementary school, since 87 percent of the underage children did not attain the achievement level of the normal age group. A sex factor appeared to be operative, also, since the underage boys made more low scores and fewer high scores than the underage girls.

The Problem

The study reported here represents an effort to obtain further evidence as to possible disadvantages encountered by children who enter first grade chronologically behind the majority of their classmates. In the area where this investigation took place, it is general policy to admit children to grade one if their sixth birthday occurs prior to the beginning of the fall term or at any time before January 1 of the next year. Members of a first grade, therefore, may range in age from five years - eight months to six years - eight months. Obviously some children have the benefit of several more months of growth than other children in their classrooms. Research in the area of child development such as the Berkeley Growth Study has provided conclusive evidence that the child's growth in most respects is still comparatively rapid during the sixth year of life. In fact, kindergarten and first-grade teachers can attest to the growth and change that are observable in individual pupils from month to month. Assuming that the experiences provided in the curriculum for a given grade level are placed there because of the readiness of the children for those experiences, one may logically conclude that learning tasks suitable for the usual six-year-old may be less appropriate for a child whose physical and intellectual status more closely resemble that of a five-year-old. Individual growth rates may contribute further to a child's immaturity.

---

* Footnotes will appear at the end of the article.

In the present study, an effort was made to test the following hypotheses:

1. Third graders who entered grade one prior to their sixth birthday do not differ significantly in reading and arithmetic achievement from children of the same grade level who entered first grade after they were six years old.
2. Children who entered first grade before their sixth birthday demonstrate as satisfactory adjustment to the school situation by the end of third grade as their older classmates.

Subjects

Twenty-nine pairs of children, selected from the third grades of five public elementary schools of upstate New York, participated in this study. Four different school systems were represented so that the effect of educational philosophy and practice would be minimized. Children were matched on sex, intelligence quotient, socioeconomic status of family, and insofar as possible, on schools attended. The variable criterion for matching was age of entrance into first grade.

In May 1960 the California Test of Mental Maturity, Primary 1957 Edition, was administered to 14 third-grade classes, from which the sample was drawn. Matching on socioeconomic status was done with the help of teachers and cumulative record data. The usual criteria, education and occupation of parents, family income, cultural advantages in the home and kind of home, were considered. No child with an obvious emotional problem was included. An effort to match the children with respect to educational history resulted in 22 pairs who had attended the same school from kindergarten through grade three.[3] The size of the sample was limited further by the variable criterion. One child in each pair had passed his sixth birthday before entering grade one, while the other attained the age of six in October, November or December following his enrollment in first grade. For the younger or underage group the mean age at entrance in grade one was five years – ten months as compared with six years – five months for the older group.

The final sample consisted of 16 pairs of boys and 13 pairs of girls. Third graders were selected because they were at the end of the primary or more informal block of the elementary school. Furthermore, three years appeared to be a reasonable length of time to allow for cancelling out, in part at least, the influences that individual growth rates might exert on early school performance.

The mean IQ scores of the underage and overage groups were respectively 116 and 115. Medians in the same order were 115.73 and 115.23.

Collection of Data

Early in June, the California Reading and Arithmetic Achievement Tests, Elementary Form CC, were administered to the 58 children. This form, rather than that for the Primary level, was used because all work included in the third grade curriculum for each school had been covered at that point. Data on adjustment were obtained by means of a scale devised specifically for this study with the help and consensus of experienced primary teachers. A survey of report and cumulative record forms suggested the list of ten traits commonly used as criteria of school adjustment. Following each trait four descriptive statements that involved implicit comparison with other children of the same age group were arranged to define four degrees or levels. The first two of these statements described behavior that was 1) clearly or 2) somewhat unfavorable, while the third and fourth items indicated evidence of the attribute to a degree 3) slightly or 4) well above that shown by the usual eight-year old.

Analysis of variance techniques were used to determine the significance of differences between groups with respect to two measures of both reading and arithmetic achievement. Numerical values of one to four were assigned to the items following each trait. Total scores and means were then computed for every trait for each group. Also determined for all traits were the percentages of children in each group whose teachers checked one of the favorable descriptive items. These data also were tested for significant differences.

Findings

Inspection of raw scores on the achievement tests shows that the overage children made consistently higher scores than their younger classmates (Table I). Analyses of variance revealed statistically significant differences between means of the age groups in all four measures of achievement. Differences were significant at the .01 level of confidence for vocabulary, comprehension and arithmetic fundamentals, and at the .05 level for arithmetic reasoning. Grade placement scores for total reading (Table II) show that more than twice as many children in the overage group made scores above 5.0 as in the younger group. In total arithmetic the difference is even greater. Two of the underage children, as compared with 11 of the older group, showed achievement above the 5.0 level.

Although differences between the sex groups were not statistically significant, the data suggest that the boys tended to find reading more difficult than the girls. Mean grade placement scores (total reading) of boys and girls in the underage group are 4.57 and 5.01 respectively; for the overage children the corresponding means are 5.31 and 5.59. Other evidence of a sex factor was provided by the lower scores of the younger boys. Three had scores below 4.0, which was the lowest score made by a younger girl. Six of the underage girls attained scores of 5.1 or higher, whereas this was true of only 3 of the younger boys. A comparison of the total groups of boys and girls shows that 44 percent of the boys and 57 percent of the girls made placement scores a year or more above actual grade level. Also of interest is the finding that six of the seven children with grade placement scores above 6.0 were girls. These findings are in line with the accumulated evidence of a relationship between the slower maturation rate of boys and the higher incidence of reading disability among them.

Teachers' Ratings of Adjustment

The ratings obtained from teachers on the ten adjustment criteria showed slightly higher means for the overage group in every instance except for responsibility where means were the same. The greatest differences were found in ratings of independence, attention span, and persistence but these were not great enough to be statistically significant. The two groups received fairly similar ratings in attitudes toward adults, social acceptance and respect for rules.

When data were analyzed for percentages rated favorably on the various adjustment criteria, the older children made a better showing than the younger on persistence, attention span, self-confidence, initiative, independence, social acceptance and social maturity. Differences on only three, however, reached or approached significance (Table III). Twenty-one, or 72 percent of the younger group as compared with 19, or 65.5 percent of the older group received favorable ratings on responsibility. Of the 10 traits listed as adjustment criteria, this was the only one on which more of the younger children were rated favorably. On seven of the ten criteria, higher percentages of the older group were judged by their teachers to be above the average third grader.

Conclusions

The data obtained in this study do not support the hypotheses formulated at the outset. Rather, they suggest that even a few months of additional growth and development may constitute an advantage when the child embarks upon his formal education. This appears to be the only factor to explain the significantly better scholastic progress made by the overage children in this study. Although differences between sex groups are not of statistical significance, they are in the expected direction and support the accumulated evidence of a sex factor that is operative in early school achievement. Teachers' ratings of adjustment, which for the most part were not significantly different, tended to be more favorable to the overage children.

To make recommendations from these findings and the data of other similar studies would mean adherence to a pattern altogether too common in educational practice, namely, the misuse of small amounts of data from the investigations of a few people. These findings, however, are adequate to justify caution in the school placement of five-year-olds. Also readily apparent is the need for further research on the problem of the underage first grader. The question raised is that of the timeliness of the usual first-grade learning tasks for children whose maturity level is not likely to equal that of children for whom these tasks were planned. More specifically, the issues are 1) whether the school functions at peak efficiency when it asks children to do school work commonly deemed appropriate for more mature children, and 2) whether such a practice serves to help or hinder the child in his efforts to live successfully and happily in the school situation.

Literature in the area of child development indicates general acceptance of two relevant principles, presumably based upon research data. One of these says, in essence, that the benefit a child derives from opportunities to acquire an understanding or skill depends to a large extent upon his level of maturity when the opportunities are provided.[4] The second principle holds that in learning something new, a child will gain as much competence from a short period of practice when he is older as he would gain from a longer period of exercise at an earlier age. If these assump- are valid, it is questionable whether the school makes the best use of its resources and of the child's time when it introduces to him curricular content which he could learn more easily and quickly a few months later.

The harmful effect of failure and frustration upon the young child also receives emphasis in the literature. Efforts to protect him from failure are reflected in administrative devices such as ability grouping, automatic promotion and the ungraded primary classroom. It is questionable, however, whether any scheme enables a child to escape the frustration that accompanies the slow and laborious learning of unreasonably difficult material. The impact of this frustration upon the child's attitude toward himself, the school and the subject in which he encounters the difficulty may be serious and lasting, even though he is spared the hurt of being labeled a failure. Nor can one disregard the fact that a major hurdle in the primary grades is that of gaining a sense of safety, security and adequacy in the school situation. Surely the child who is under pressure to master learning tasks beyond his level of maturity faces compounded adjustment problems.

In conclusion, it appears logical to account for the lower achievement of the underage group in this study by assuming that they had been given learning experiences before they reached their peak performance in coping with those tasks. Findings support the results of previous studies in

TABLE I

ACHIEVEMENT TEST SCORES
QUARTILES FOR UNDERAGE AND OVERAGE GROUPS

Subtest	First Quartile		Median		Third Quartile	
	Underage Group	Overage Group	Underage Group	Overage Group	Underage Group	Overage Group
Reading Vocabulary	56.75	64.38	61.75	72.06	69.63	75.06
Reading Comprehension	16.75	22.63	20.75	25.07	24.60	28.67
Arithmetic Fundamentals	13.0	17.56	17.3	19.38	19.43	22.29
Arithmetic Reasoning	17.75	19.55	22.75	24.25	25.71	29.08

TABLE II

TOTAL READING AND ARITHMETIC GRADE PLACEMENT SCORES FOR UNDERAGE AND OVERAGE GROUPS

Scores	Reading		Arithmetic	
	Underage Group	Overage Group	Underage Group	Overage Group
3.1-3.5	1		1	
3.6-4.0	3	1	3	2
4.1-4.5	9	4	8	5
4.6-5.0	7	4	15	11
5.1-5.5	4	8	2	9
5.6-6.0	4	6		2
6.1-6.5		3		
6.6-7.0	1	3		
TOTAL	29	29	29	29

TABLE III

PERCENTAGES OF CHILDREN IN EACH GROUP WHO WERE RATED FAVORABLY ON TRAITS COMMONLY CONSIDERED CRITERIA OF SCHOOL ADJUSTMENT

Trait	Percentage		
	Underage Group	Overage Group	Difference
Persistence	59	79	20
Attention span	21	52	31**
Self-confidence	41	52	11
Responsibility	72	65.5	6.5
Initiative	24	38	14
Independence	38	65.5	27.5**
Respect for rules	64	64	—
Social acceptance	64	72	8
Social maturity	38	62	24*
Attitude toward adults	64	64	—

**Significant at the .05 level

* Approaches significance

questioning the soundness of placing children in first grade prior to their sixth birthday. Attention is called also to the lag in translating into practice the accumulated research data on sex differences in readiness for initial reading instruction.

An answer to the problem of the underage first grader appears to be especially important at the present time because of current emphasis on gifted children and the possible reluctance of parents to "hold them back." A major criterion for identifying a gifted child appears to be his performance on an intelligence test. Studies such as the one reported here suggest that factors not reflected in an IQ score have a substantial effect upon a child's early school success.

Further research with large numbers of children would furnish conclusive evidence. Until this is available, it may be advisable to defer first-grade enrollment for some children who normally would enter underage. With a good foundation in the early grades, acceleration quite conceivably could be accomplished at a later level with a minimum of risk.

FOOTNOTES

1. Baer, Clyde J. "The School Progress and Adjustment of Underage and Overage Students," Journal of Educational Psychology XLIX (1958), pp. 17-19.

2. Carter, Lowell B. "The Effect of Early School Entrance on the Scholastic Achievement of Elementary School Children in the Austin Public School," Journal of Educational Research, L (Sept. 1956), pp. 91-103.

3. Note – Only one child was included who had not attended kindergarten. She had been admitted to a non-public school which lacked a kindergarten because of her superior performance on an individual intelligence test.

4. Jersild, Arthur T., Child Psychology, (New York: Prentice-Hall, 1960), p. 21.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 8, April 1963)

A Comparison of Two Methods for Determining Difficulty In a Multiple Choice Test *

MARGERY V. HAWES
Wheat Ridge, Colorado
GERALD C. HELMSTADTER
Arizona State University

TRADITIONALLY, item difficulty has been measured by determining the percentage of examinees in a group who correctly answered the question. Although most persons have used this classical method without question, some dissatisfaction with it has been expressed. See, for example, Davis (1951). The major limitation of using the proportion who get the item correct is that is does not reveal differences among individuals with respect to the degree of difficulty they have on each single item. Thus, for one person, an item may be troublesome because he has difficulty in making a choice between only two of the alternatives presented; for another, all of the alternatives of the same item may appear as reasonable answers.

The rationale for ignoring such differences in the usual measure of difficulty assumes that, when the size of the group is large enough, the laws of chance will operate in such a way that the expected value for the number of persons who will get the item correct will occur. Even so, it is possible to have items with identical proportions of persons who get the correct response which, intuitively, are not of the same difficulty. Consider, for example, the hypothetical situation for the three 5-choice items represented in Table I. While the expected proportion of persons who mark the item correctly is identical for all three, the breakdown, showing the different ways in which that proportion was achieved, suggests that the items are definitely not of equal difficulty. Item B, where 15% of the group could eliminate none of the alternatives and an additional 16% only one alternative, would appear to be the most difficult. Item C, where every person was able to eliminate at least two alternatives and where over half of the group eliminated three of the five choices, would seem to be the easiest.

In spite of this limitation of the classical measure, no one seems to have suggested any other way of determining item difficulty. The problem of this study, therefore, was that of devising and evaluating a new measure of item difficulty which considers the response of each individual to each item.

The Measures of Difficulty

One logical basis for obtaining a measure of the difficulty of each item of a test for each of the examinees who takes the test, appears to be a count of the number of responses required for each individual to obtain the correct answer.[2] Although this approach does not distinguish those items a respondent easily gets right from those he answers correctly only with some thought, it does permit an indication of the degree of difficulty a person has with those he gets wrong. If an individual makes five responses before he gets the correct answer, the chances are it is a more difficult item for him than one which requires only two selections.

To obtain a single descriptive index representing the difficulty of a particular item for an analysis group, it is necessary to consider the distribution of the number of responses made to the item by the members of the group. Such a distribution for each of the three hypothetical items discussed previously is presented in Table II. It can readily be seen that this distribution will be highly skewed for easy items, with the greatest number of persons requiring only one or two choices. As the items become more difficult the distribution will become more and more rectangular, the extreme case occuring when everyone in the group makes a pure guess among all alternatives, and there is an equal probability that the correct response will be found on any one of the tries.

Obviously, the desired index of difficulty for each item should be some measure of central tendency of this distribution. Although there are many situations in which the median is appropriate when the distribution is skewed, it is the very fact that values occur throughout the possible range that intuitively leads us to say that one item is more difficult than another. Thus, in this instance, the mean,

---

*Footnotes will be found at end of article.

which is influenced by the magnitude rather than the rank position of the observations is most appropriate.

Because this new measure permits an indication of the difficulty of each item for each individual as well as providing an overall index for each item, it will be called individual item difficulty (IID). For contrast, in this article, the classical procedure of computing the proportion of the group that answered the item correctly on the first try will be called group item difficulty (GID).

Procedure

The Test Medium

To compare the two ways of measuring item difficulty, it was necessary to develop a method by which measures of both GID and IID might be obtained from one presentation of a test. Also, it was essential that the test used contain items which covered the full range of difficulty and that each item had a sufficiently large number of plausible distractors to permit as wide a selection of alternatives as possible by the individual. Otherwise, the resulting restriction in range of the individual item difficulty measure might tend to obscure whatever relationship between the two measures might exist.

One type of test which seemed amenable to the above requisites was a vocabulary test. Thus, eighty items were constructed, each having a stem word, one synonym and fourteen plausible distractors. These items were formed into a pre-test and administered to twenty-nine high school juniors and seniors, twenty-six university freshmen registered for a beginning English class, and ten faculty members.

For this testing, parentheses were placed before each of the fifteen alternatives. Each examinee was asked to first place the number "1" in the parentheses before the alternative which he thought represented the meaning of the stem word and then, to help check the adequacy of the distractors, to read through the alternatives a second time. This time the subjects were to continue ranking the alternatives according to their personal judgment as to the value of these distractors for use in a later test.

From the information obtained on this pre-testing, it was possible to construct a testing instrument consisting of sixty items of widely varying difficulty arranged in two parallel parts of thirty items each.

The Response Sheet

In addition to the special test that had to be constructed, it was necessary to develop a response device that would record the number of tries required by each individual to answer each item correctly. For this purpose a special response sheet was constructed which incorporated ideas suggested by the cardboard punch described by Angell and Troyer (1948) and by the tab item as developed by Glaser et al.(1954).

Fifteen circular, perforated positions for each of ten items were punched in an 8 1/2 × 11 inch piece of chipboard. Each of the tabs thus formed was numbered to correspond to one of the responses for each item. Working from the reverse side of the chipboard, a large red dot was then painted on the underside of each tab which corresponded to a correct answer. Finally, the underside of the chipboard was covered with a heavy piece of manila paper, thus preventing the examinee from finding the answer by turning the response sheet over.

The Test Administration

For the test administration, each subject was provided a heavy pin with which to lift the response tabs. The test directions instructed the examinees to look at the reverse side of the tab each time they made a response, and to continue making choices for each item until the correct one, indicated by a red dot, was selected. The number of tabs lifted thus provided a measure of the difficulty of each item for each individual. Items for which only one tab was removed were counted as correct for the classical or group item difficulty measure.

To get the data for the analysis, the test was administered to four sections of university students registered for general psychology classes during the winter and spring quarters of the 1958-1959 school year. Thirty of the total test items were administered on each of two days during the regular 50 minute class period. Altogether, 244 subjects were tested. However, the results of 13 of these were unusable: 11 failed to follow instructions, one overlooked two items, and one failed to cooperate on the second half of the test. Thus, the results are based on data obtained from a final sample of 231 individuals.

Results and Discussion

Total Test Scores

With a possible score of 60, the total test scores for the 231 subjects ranged from 13 to 44. The mean was 24.9 and the standard deviation was 5.7. The distribution was slightly skewed, indicating that the test, over-all, was a bit too difficult for the group used.

As mentioned in the previous section, the 60 items, roughly matched item for item with respect to difficulty, were organized into two separate parts so that thirty words could be administered on each of two days. Using the two separately administered parts, a Guttman reliability (Anastasi 1954, p. 109) was computed and found to be .48. Because this unusually low reliability might have resulted in part from an unequal division of the test into two halves, it was recomputed using an odd-even split, and found to be .58 One possible explanation for this disappointingly low reliability (especially for a vocabulary test) is that because the items were somewhat difficult, an unusual amount of guessing occurred. Fortunately, as will be seen later, this did not seem to affect either measure of difficulty.

The Measures of Difficulty

Computation of the classical measure of difficulty for each of the sixty items indicated that a range of difficulties from .06 to .97 had been achieved. As suggested by the distribution of the total test scores, there was a heavy concentration at the lower end of the scale, approximately 50% of the items falling below a difficulty index of .35. The means of the distribution of the number of responses required to get the answer correct ranged from 1.12 to 7.03. One-third of the items had an index of five or larger and approximately one-third a mean of less than 2.

To estimate the reliability (sample to sample consistency) of each of the two measures of item difficulty, the 231 subjects were divided into two groups matched according to the total scores obtained on the test. Both the individual item difficulty and group item difficulty were then determined for each of the items, first using the results obtained on group A and then the results from group B. Next, the correlation over items and between groups was computed. This was found to be .96 for group item difficulty and .98 for individual item difficulty.

It is especially interesting that the reliability of the measures of difficulty are very high in spite of the low reliability of the test as such. This suggests that, even though an individual's score might vary considerably when given different forms of the same test, the measures of the difficulty of the items in the test are stable over groups roughly matched for over-all ability in the function mentioned. Also, it is apparent from the similarity of the two reliabilities, that one measure cannot be said to be more adequate than the other as far as stability from group to group is concerned.

The Relationship Between Measures of Difficulty

To determine whether the two measures of difficulty do produce essentially similar results, the product-moment correlation between the two measures over items was computed and a test of significance (cf. McNemar, 1958) of whether the correlation between the two measures was farther below unity than attributable to measurement error was made. Although the correlation was found to be .94 (because of the way in which IID was measred. this negative correlation indicates the two measures are in agreement) the item by measure interaction was significant. Thus, it could not be concluded that the two indexes are identical except for errors of measurement.

When the scatter plot between GID and IID was made, however, it was apparent that the relationship between the two measures of difficulty was curvilinear. Thus, the product-moment correlation as computed was an underestimate of the degree of relationship between the two variables.

To obtain a better indication of the agreement between the two different measures of difficulty, the correlation ratio, eta was obtained. However, since eta is a regression coefficient, it was first necessary to decide whether to use the regression of the GID measure on the IID measure, or the regression of the IID on the GID measure Both because individual item difficulty provides a measure of the difficulty of each item for each person, and because it was intuitively more appropriate as a group index, it seemed reasonable to consider the classical procedure as only an easily obtained practical estimate of the former. Thus, to answer the question of the extent to which IID can be predicted from GID, the regression of IID on GID was found. Using 10 class intervals of length 0.1, eta turned out to be .96.

It is thus apparent that while the new measure is not identical with the traditional one, and provides additional information which may be exceedingly useful for individual teaching and diagnostic use, for most practical test development procedures, it will make little difference which index is used.

Summary and Conclusion

The major purpose of this study was that of devising and evaluating a measure of item difficulty which considered the individual's different responses to the item. To accomplish this, a multiple-choice vocabulary test was constructed consisting of sixty items, each with fifteen alternatives. A special tab response sheet was devised so that the examinees would know whether they had the right answer when they made their choice. Thus, instead of following the usual directions, the examinees were asked to continue responding to each item until they had selected the correct answer.

The test was administered to 231 subjects registered for general psychology classes during the winter and spring terms in the 1958-59 school year at Colorado State University.

From the results of this testing, two measures of item difficulty were obtained. First, the classical procedure of computing the proportion of the group that answered the item correctly on the first try was followed. Then, the individual item difficulty was obtained. This was accomplished by counting the number of responses made to each item by each person and then, for each item computing the arithmetical mean of the number of responses made by all individuals.

Reliabilities of the measures were determined by dividing the total sample into two groups matched for over-all test scores, and by computing the correlation between similar measures of difficulty, over items. Reliability was .96 for the group item difficulty measure and .98 for the individual item difficulty measure.

Using the total sample, a scatter diagram showing the relationship between group item difficulty and individual item difficulty over items was made. Since the relationship between the measures of difficulty was curvilinear, the correlation ratio, eta, was used to describe the degree of similarity and found to be .96. A test of significance of whether the correlation between the two indexes was farther below unity than could be attributabe to measurement error indicated the measures were not identical.

On the basis of these results it was concluded that counting the number of responses required to obtain the correct answer would provide a useful index for obtaining the difficulty for an individual but that for practical purposes in test development, it would make little difference which of the two indexes was used.

FOOTNOTES

1. This study was done in partial fulfillment of the requirements for a Master's degree at Colorado State University and was supported in part by the Colorado State University Research Foundation.
2. This concept of difficulty was suggested to the junior author when he was a graduate student at Iowa State University by Dr. Charles O. Neidt, now at the University of Nebraska.

REFERENCES

1. Anastasi, Anne. Psychological Testing, New York: Macmillan Company, 1954.
2. Angell G. and Troyer, M. "A New Self-Scoring Test Device For Improving Instruction." School and Society, LXVII (1948), 84-5.
3. Davis, F. "Item Selection Techniques" in E. F. Lindquist (Ed.) Educational Measurement, Washington, D.C.: American Council of Education, 1951.
4. Glaser, R.; Damerin, Dora E.; and Gardener, F. "The Tab Item: The Technique for the Measurement of Proficiency in Diagnostic Problem Solving Tasks" Educational and Psychological Measurement, XIV (1954), 283-93.
5. McNemar, Q. "Attenuation and Interaction" Psychometrika, XXIII (1958), 259-65.

TABLE I

HYPOTHETICAL RESPONSES TO THREE ITEMS HAVING IDENTICAL EXPECTED DIFFICULTY INDEXES

No. of Choices Among Which S's Guessed	Item A		Item B		Item C	
	No. of Persons	Expected No. Marking Correctly	No. of Persons	Expected No. Marking Correctly	No. of Persons	Expected No. Marking Correctly
1	20	20	22	22	15	15
2	30	15	32	16	40	20
3	30	10	15	5	45	15
4	20	5	16	4	0	0
5	0	0	15	3	0	0
Total	100	50	100	50	100	50
Expected Proportion Correct	.5		.5		.5	

TABLE II

EXPECTED FREQUENCY DISTRIBUTIONS OF THE NUMBER OF RESPONSES REQUIRED TO CORRECTLY ANSWER THE ITEM

X	Item A						Item B						Item C					
Number of Responses Required	Number of Choices Among Which S's Guessed						Number of Choices Among Which S's Guessed						Number of Choices Among Which S's Guessed					
	1	2	3	4	5	Total A	1	2	3	4	5	Total B	1	2	3	4	5	Total C
1	20	15	10	5	0	50	22	16	5	4	3	50	15	20	15	0	0	50
2	0	15	10	5	0	30	0	16	5	4	3	28	0	20	15	0	0	35
3	0	0	10	5	0	15	0	0	5	4	3	12	0	0	15	0	0	15
4	0	0	0	5	0	5	0	0	0	4	3	7	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0	0	3	3	0	0	0	0	0	0
Total	20	30	30	20	0	100	22	32	15	16	15	100	15	40	45	0	0	100
$\Sigma f_i X$			1.75						1.85						1.65			
IID			1.75						1.85						1.65			

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 8, April 1963)

Educators' Attitudes Toward Educational Practices

ROBERT S. MILLER
Washington State University

DURING THE PAST several years, the philosophy and practice of American public education have come under fire from various quarters. "Oceans of piffle," "intellectual pablum," and "twaddle" are some of the terms applied. Some of the critics have insisted that many of their pronouncements found willing listeners within the ranks of professional educators, and have alluded to supporting testimony.

The present study was undertaken to assess attitudes held by various sub-groups of professional educators in regard to the previously mentioned controversy. The object has been to test specifically the differences of attitude within the profession between various subgroups, as described later.

Kerlinger (2) has demonstrated by another approach and for a different purpose that two basic attitudes toward educational philosophy exist, named "progressivism" and "traditionalism." In the present study they are called "child centeredness" and "subject centeredness." No attempt has been made to test whether the two studies are sampling the same latent variables, but it seems highly probable.

Another attitude studied here was that subsumed under the title "school community cooperation." Some of the critics have said that control of the schools has passed from the community to the professional educators, and respondents for this study were asked to react to items concerning this issue.

These objects of attitude constitute one dimension of the attitudes studied, with the second dimension being that of the character of being favorable or unfavorable. There are six attitudes to be studied, "subject centered favorable," "subject centered unfavorable," "child centered favorable," "child centered unfavorable," "school-community cooperation favorable," and "school-community cooperation unfavorable." More sophisticated labels could be chosen to replace these unwieldy nes, but these have the virtue of being self explanatory.

If there is an objection that such contrasting pairs of attitudes as "child centered favorable" and "subject centered unfavorable" appear to be only two sides of the same coin, the answer lies in the validation procedure for the questionnaire.

The object of the study should now be stated more precisely as being a test of whether various subgroups of professional educators may respond similarly in agreeing or disagreeing with these named attitudes, as they are exemplified in the questionnaire.

The Questionnaire and The Items

Items or single statements of attitude representing the six attitudes were drawn from the publications of various critics and from the publications of the defenders of public school procedures. Over seven hundred items were originally collected. Four professional psychologists and educators, working separately, judged the attitude category into which each item should be placed. All four judges agreed on the proper categorization of each of 140 items. This unanimous judgment made it unnecessary to compute scaling or validity statistics, since building attitude scales as such was not the object, but rather the identification of items really representing the attitudes. These attitudes, as described here, are purely nominal scales.

From the pool of 140 items a random sample of ten for each attitude was drawn. This made a questionnaire of sixty items.

The Professional Subgroups

Two points were kept in mind in specifying these subgroups, first, the kind of job held, and second, geographical location of the respondent. It was hypothesized that differences in attitude might exist between teachers of differing kinds of subjects, and that some factor subsumed under the facto of location might be responsible for some differences in attitude.

Teachers were categorized as Academic and Non-Academic, the assumption being that the teaching of such subjects as mathematics, language arts, social sciences, natural science, and the like is subject to rather different procedures than is the teaching of such subjects as agriculture, shop, driver training, music groups, or commercial subjects. Non-academic subjects require a closer and more individualized kind of instruction, which suggests the possibility of different attitudes about teaching. It was possible to categorize accurately every individual who responded to the questionnaire, on the basis of college preparation and present teaching assignment.

Two other subgroups included for study were school administrators and subject matter department heads or supervisors. These two groups tend to be removed from the actual classroom situation in greater or less degree, and because of the duties of their jobs might be thought to differ in attitudes about teaching from each other and from teachers. They may also have an influence on the attitudes of teachers.

The States of Kansas and California were chosen to represent the variable of geographic location. The writer had some familiarity with them, and it has been suggested that there are differing attitudes toward education in the two states. However, this variable may not take sufficiently into account the westward migration of teachers recently.

School directories of the two states provided names of administrative officers, some of whom in turn provided faculty lists. Seventy-four school districts in California and fifty in Kansas were asked to supply such lists. Of these, thirteen (18%) of the California administrators supplied usable faculty lists, and twenty-two (45%) of the Kansas administrators complied. A few administrators replied that supplying such lists was contrary to local policy. Certain California administrators replied that their state association of secondary school principals must pass judgment on such samplings of opinion, but data were already being collected from California before this came to light. However, it may explain in part the rather small proportion of responses from California administrators.

On each list thus obtained, every second teacher, department head or supervisor, and administrator on the secondary level was addressed an invitation to participate. This was in December, 1957. A return post card was included, upon which the individual could indicate willingness to participate. A total of 640 educators received the initial invitation, and 392 (61%) returned the post card. Of those who replied, 342 (87%) agreed to respond to the questionnaire. Usable questionnaires returned numbered 240. This was 37.5 percent of the original group canvassed.

A question about this sample may be raised. Any sampling error occurring because of the nature of the group of administrators supplying faculty lists has its root in the way school districts were listed in the directories. Common sense suggests that the findings of this study could be generalized to the whole profession in the two states, but the obvious shortcomings of the sample prevent giving this generalization statistical support.

Statistical Procedures

The Q-methodology of William Stephenson (3) has intriguing possibilities for providing estimates of the attitudinal private world of an individual, since it provides for assigning one score from a range of scores, as a response to a single item, and the subsequent possibility of treating a distribution of scores for each item. Stephenson's methods are well known and need not be detailed here. Cronbach (1) describes them as a "bundle of loosely related devices which can be separately adopted or rejected." Two devices from this bundle were utilized in the present study, obtaining responses or scores by a variation of card sorting using scores printed on the questionnaire for the respondent's use, and selecting questionnaire items according to a structured or "balanced block" design. Higher scores indicated agreement with each statement, and lower scores disagreement.

Stephenson suggests both correlational and analysis of variance procedures for treating Q-sort data. In the present study neither was usable because the distributions obtained were extremely skewed in many instances, and completely J-shaped in several. Further, the obtained data are from purely nominal scales, and are not ordered or linear or continuous.

Hence the statistic chosen was chi square, since it is concerned with nothing but the distribution of expected frequencies around observed frequencies. To accommodate the data from the fifteen point range available to respondents to treatment by chi square, the range was divided into five three-point segments, corresponding to strong disagreement for scores one through three, mild disagreement for four through six, neutral or undecided for seven through nine, mild agreement for ten through twelve, and strong agreement for thirteen through fifteen.

Findings

Four analyses of the data were made. The first involved a comparison of the responses of the six subgroups to each of the sixty items. For each item a chi square solution was obtained, with six columns for the six subgroups and five rows for the five degrees of response. Scores from all respondents were included. Of the sixty items, six differentiated between the subgroups at the .05 level of confidence or better. A formula for computing chi square values in tables of this size was taken from Walker and Lev (4) and used exclusively in the study.

In the questionnaire, each attitude is represented by ten items. In the second treatment of the data, responses to each ten-item attitude were accumulated and the responses of the six professional subgroups compared to these attitudes. Table I shows the probability levels obtained by chi square comparisons. The columns in the chi square solution were again the six subgroups and the five rows the degrees of response.

TABLE I

PROBABILITY LEVELS OBTAINED BY CHI SQUARE COMPARING RESPONSES FROM ALL SUBGROUPS TO THE SIX ATTITUDES

Attitudes		Probability Levels
IA	Child Centered Favorable	.001
IB	Child Centered Unfavorable	.02
IIA	Subject Centered Favorable	.05
IIB	Subject Centered Unfavorable	.02
IIIA	School-Community Cooperation Favorable	.02
IIIB	School-Community Cooperation Unfavorable	.001

Such levels of confidence appear to indicate true independence between the subgroups, or differences in response to the attitude stimuli. These differences, appearing as they do in all six comparisons, strongly suggest that there are indeed differences in attitude between these subgroups. This finding does not support the contention that many professional educators agree with the critics of edcation. Inspection of the tabulated responses to questionnaires, not presented here, indicate only that the subgroups differed in relation to each other, but no trend to agree with the critics is discernible.

The comparison just described involved all six professional subgroups in a single comparison on each attitude, with the results noted. It is possible, however, that such a comparison might mask differences between individual pairs of subgroups. Accordingly, each possible pairing of subgroups was compared in regard to their responses to each attitude, by the same chi square procedure. Since this was a detailed test of the major point of this study, the levels of confidence obtained by these comparisons are presented in full in Table II. It is seen there that of the ninety comparisons, twenty-five pairings of subgroups show significant differences. This lends further and very convincing support to the general finding that professional subgroups of educators differ in their ideas about educational philosophy. It is also clear that these differences occur in only a significant minority of the comparisons.

A final point of interest concerned whether the professional subgroups distinguished between the favorable and unfavorable attitudes. Table III shows levels of confidence obtained. That these individuals responded differently, or differentiated between favorable and unfavorable attitudes, is unmistakable, even though in one instance the difference did not reach usually accepted levels of significance.

Summary

This has been a study to determine whether there are differences of attitude between defined subgroups of professional educators in regard to the controversy over educational procedures. Critics have urged a return to a more subject-centered system, as against a permissive and child-centered system.

A questionnaire was constructed exemplifying six attitudes, and respondents were asked to indicate agreement or disagreement with statements representing these attitudes. The responses of six subgroups were compared by chi square.

Individual items did not differentiate well between subgroups, but each ten-item attitude differentiated significantly both when responses from all six subgroups were compared together, and when subgroups were compared by pairs. These differences were demonstrated in a significant minority of comparisons. Finally, it was demonstrated that the responding educators distinguished clearly between favorable and unfavorable attitudes.

REFERENCES

1. Cronbach, L. J. "A Review of Stephenson's Q-Technique and Methodology." Psychometrika XIX (December 1954), 327-30.
2. Kerlinger, Fred N. "Progressivism and Traditionalism: Basic Educational Attitudes." School Review LXVI (March 1958), 80-92.
3. Stephenson, William. The Study of Behavior: Q-Technique and Its Methodology (Chicago: University of Chicago Press, 1954).
4. Walker, Helen M., and Joseph Lev. Statistical Inference (New York: Henry Holt and Co., 1953.)

TABLE II

PROBABILITY LEVELS OBTAINED BY CHI SQUARE COMPARING RESPONSES OF EACH POSSIBLE PAIR OF SUBGROUPS TO EACH ATTITUDE

Pairs of Subgroups	Attitudes IA	IB	IIA	IIB	IIIA	IIIB
Kansas Academic Teachers California Academic Teachers	.001	.20	.20	.001	.01	.20
Kansas Academic Teachers Kansas Non-Academic Teachers	.80	.70	.20	.70	.01	.10
Kansas Academic Teachers California Non-Academic Teachers	.50	.70	.20	.70	.50	.20
Kansas Academic Teachers Administrators	.20	.50	.50	.10	.80	.70
Kansas Academic Teachers Department Heads	.02	.05	.01	.05	.50	.01
Kansas Non-Academic Teachers California Academic Teachers	.01	.01	.70	.01	.98	.01
Kansas Non-Academic Teachers California Non-Academic Teachers	.80	.50	.10	.70	.70	.01
Kansas Non-Academic Teachers Department Heads	.05	.01	.05	.05	.50	.001
Kansas Non-Academic Teachers Administrators	.20	.30	.50	.50	.05	.01
California Academic Teachers California Non-Academic Teachers	.20	.50	.50	.10	.80	.70
California Academic Teachers Department Heads	.01	.30	.30	.20	.50	.10
California Academic Teachers Administrators	.001	.05	.70	.30	.02	.90
California Non-Academic Teachers Department Heads	.30	.20	.70	.10	.99	.70
California Non-Academic Teachers Administrators	.20	.30	.20	.90	.30	.90

TABLE III

PROBABILITY LEVELS OBTAINED BY CHI SQUARE COMPARING RESPONSES TO FAVORABLE AND UNFAVORABLE ASPECTS OF ATTITUDES FOR EACH SUBGROUP INDIVIDUALLY

Subgroup	Pairs of Attitudes IA/IB	IIA/IIB	IIIA/IIIB
Kansas Academic Teachers	.001	.001	.001
Kansas Non-Academic Teachers	.001	.20	.05
California Academic Teachers	.001	.001	.001
California Non-Academic Teachers	.001	.01	.001
Department Heads	.001	.001	.001
Administrators	.001	.02	.001

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 8, April 1963)

Ability of Poor Readers to Discriminate Paired Associates Under Differing Conditions of Confirmation*

WAYNE OTTO
University of Oregon

THE PRESENT STUDY is an attempt to get at reasons for certain relationships that were demonstrated in an earlier study (Otto, 1961a, 1961b) of paired associates learning by good, average, and poor readers. Very briefly stated, the earlier study showed that 1) good, average, and poor readers differ in the number of trials required to learn a list of paired associates, and 2) mode of reinforcement (method of confirmation of the correct response) and grade placement-chronological age of the subjects significantly affects the number of acquisition trials required.

A possible reason for the difference in learning efficiency exhibited by the groups of good, average and poor readers appears to be that poor readers may have difficulty in differentiating between the stimulus figures and between the response trigrams that make up the paired associates list. Furthermore, for poor readers the act of associating an enunciated syllable with a printed trigram may have amounted to an additional associative learning task within the ostensible learning task. The present study, then, was designed to investigate these notions.

Method

Subjects

Thirty poor readers, who were enrolled in the 1961 University of Wisconsin Summer Reading Clinic at the time of the study, served as subjects. Ten subjects were assigned to each of three modes of reinforcement.

In all cases, reading achievement was substantially below expectancy according to both standardized tests and teacher judgment. The poor readers of the present study were, then, comparable to the poor readers of the earlier study in terms of capacity/achievement. However, while in the earlier study all subjects met criteria of an IQ (group test) score in the 95-110 range and placement in grade 2, 4, or 6, the present subjects had IQ (individual test) scores ranging from 92 to 129 and grade placement in grades 4 to 7 as of September, 1961.

Task and Reinforcement

The task and modes of reinforcement followed that of the earlier study. The paired associates list consisted of five familiar, geometric forms (diamond, circle, triangle, star, square) paired with five low association value CVC trigrams (fep, miv, wuc, yad, gox). The pairs were presented by a Stone memory drum at 4 second intervals. Confirmation was provided by the experimenter, who articulated the proper trigram response and presented the printed trigram.

Three modes of reinforcement (confirmation) were employed: auditory, visual-auditory, and kinesthetic-visual-auditory. For the sake of brevity, these are called simply auditory, visual, and kinesthetic. Auditory reinforcement consisted simply of the experimenter's articulation of the trigram; the subject did not see the printed trigram at any time. Visual reinforcement combined articulation of the trigram with presentation of the printed trigram. Subjects receiving kinesthetic reinforcement traced the letters of the trigram while the trigram was named by the experimenter.

As in the earlier study, total acquisition trials was the sum of trials required for one correct anticipation of the list with serial presentation plus one correct anticipation of the list with scrambled presentation of the pairs.

Procedure

Since one of the primary purposes of the present study was to determine whether poor readers had difficulty in discriminating the forms and trigrams of the paired associates list, the procedure differed somewhat from the earlier study.

Discrimination Training. Before beginning his acquisition trials, each subject's ability to discriminate between the trigrams and forms employed in the paired associates list was tested.

---

*Given as a paper at the annual meeting of the American Educational Research Association, February 19-21,

Each subject was shown each form individually and asked to give whatever name he had for it. As pointed out above, the forms were familiar and each subject did, indeed, have readily available a name for each one. The subject was then shown all of the forms at once and asked to name each one.

Ability to discriminate trigrams was also tested. Subjects to receive auditory reinforcement were tested for their auditory discrimination of the several trigram names: the experimenter would state a key trigram name and then a series of three trigram names, one of which was the key; the subject was then asked to decide whether one, two, or three was the key. Subjects to receive visual or kinesthetic reinforcement were first shown the printed trigrams as the experimenter pronounced them; they were then asked to give the trigram names as the experimenter presented the printed trigrams.

Subjects were given as many trials as they needed to reach the criterion of two times through the list (of forms or trigrams) without any miscalls.

Acquisition. After the discrimination training, the subjects learned the paired associates list to the criterion.

Associations. Finally, the experimenter asked for associations evoked by each form and each trigram and, when appropriate, quizzed the subjects as to why a particular pair was associated either very quickly or only after a relatively great number of trials. It was felt that analysis of such information might yield clues that could be helpful in interpreting results. For example, possible confusion arising from poor discrimination--despite the pre-training--could be uncovered.

Design. The three modes of reinforcement groups were matched by chronological age, since the earlier study showed grade placement-chronological age to significantly influence performance on the present task.

Since IQ level had been restricted in the earlier study, but not in the present study, the possible effect of the wider range of IQ scores upon total acquisition scores needed to be considered. Consequently, the design provided for matched subjects across rows with IQ and total acquisition trials as covariates. (Edwards, 1950).

Results and Discussion

A comparison of mean IQ scores of three groups showed no significant differences. The three groups were matched by chronological age.

Discrimination training. The criterion was two times through the list (form or trigram) with no miscalls. Two trials, therefore, was the least possible number of trials.

In no case did a subject take more than two trials with the list of forms. That is, in all cases the subjects were able to correctly differentiate the forms without any errors at all.

The trigram list presented some difficulty when the printed trigrams were the stimuli. When the trigram names only were presented to the auditory reinforcement group, no errors were made. When the printed trigrams were presented to the visual reinforcement group and to the kinesthetic reinforcement group, five subjects made errors. Three of these subjects required three trials and two subjects required four trials; thus, errors were made in trials one and/or two.

None of the errors, however, were total miscalls (i. e. wuc for gox). Rather, the errors were lack of response or incorrect pronunciation of the trigram. It appears, then, that the errors were due not to lack of discrimination but to failure to associate the correct pronunciation with the trigram.

It would seem reasonable to conclude that the inferior performance of poor readers demonstrated in the earlier study cannot be attributed to lack of discrimination of the forms and trigrams making up the paired associates.

The fact that five subjects had difficulty with the recognition-enunciation of the printed trigrams, however, seems to lend support to the notion, stated above, that the poor readers may have been faced with the task of learning trigram names. The extent to which the need for additional learning of this kind might have contributed to the inferior performance of the poor readers will be examined in the light of additional data.

Acquisition. It seems reasonable to hypothesize that if associating trigram and trigram names constitutes an additional learning task that depresses performance of poor readers, then poor readers receiving auditory confirmation should perform better than poor readers receiving visual or kinesthetic confirmation since the auditory confirmation requires no trigram-trigram name association.

One way of testing this hypothesis is to take a closer post hoc look at data from the earlier study, while conceding the dangers of such post hoc looks. Table I shows the mean acquisition trials of poor readers by grade level and by mode of reinforcement. The data show that the subjects receiving auditory reinforcement tended to take more trials rather than fewer, thus directly contradicting the hypothesis.

Data from the present study tend to corroborate this contradiction. Table II shows the mean acquisition trials of the poor readers in the present study by mode of reinforcement. Again, the subjects receiving auditory reinforcement–that is, no presentation of the printed trigrams–did least well.

Thus, it appears that poor readers are not so confused by the printed trigram that their performance with the present task is significantly depressed.

To the contrary, the introduction of the printed trigrams appears to enhance performance for poor readers, just as it does for average and good readers.

Statistically, the differences in mean acquisition trials shown in Table II are not significant. (F = 2.0; 3.55 required for significance at the .05 level with 2 and 18 degrees of freedom). However, the stated hypothesis would require significantly better performance by the auditory reinforcement group in order to be acceptable.

The lack of significant differences between modes of reinforcement does not agree with the significant differences demonstrated in the earlier study. It will be recalled, however, that while the present study placed no restriction upon the IQ range of subjects, the earlier study restricted IQ scores to the 95-110 range. Therefore, the design of the present study provided for the study of acquisition scores and IQ scores as covariates.

The acquisition scores, however, appear to have been uninfluenced by IQ scores. It has been pointed out that a comparison of mean IQ scores of the three modes of reinforcement groups showed no significant differences. (F = .31; 3.55 required for significance at the .05 level with 2 and 18 degrees of freedom). It was found, in fact that the correlation coefficient for IQ scores and acquisition scores was .0321. This apparent lack of relationship raises more questions than it answers.

Two conclusions seem warranted by the acquisition data. 1) The printed trigram does not appear to be a serious source of confusion for the poor readers studied. It seems necessary, therefore, to go beyond this lack of discrimination-confusion idea for an explanation of the finding of the earlier study that poor readers required more acquisition trials than did average or good readers. 2) The virtual lack of correlation between IQ and acquisition scores in the present study warrants further investigation. It is suggested that addition of IQ as a manipulated variable to the basic design of the earlier study (Otto, 1961) would be a useful way of getting at possible interactions that might help to explain the lack of correlation demonstrated in the present study.

It should be noted here that inspection of Table I and Table II reveals that the poor readers of the present study took fewer acquisition trials than did the poor readers of the previous study. That the criteria for selecting the two populations differed has been pointed out. Other possible causes can be pointed out. For example, while the discrimination training of the present subjects appeared to reveal no specific difficulty, the familiarization process alone could be expected to facilitate performance. Also, the present subjects were enrolled at the summer reading clinic; it is quite possible that their performance was enhanced by the fact that they were participating in the various clinic experiences. Any of these possible causes offer interesting possibilities for further study.

Associations. The experimenter asked each subject for associations evoked by each figure and each trigram after the acquisition trials; and, when applicable, the subject was asked if he could say why a particular form-trigram took relatively few or many trials. This was done with a view toward uncovering confusion arising from lack of discrimination of the forms and trigrams or similar phenomena.

In general, analysis of the associations given revealed few cases where associations evoked in a particular subject would appear to have a detrimental effect on acquisition. In five cases, however, there was some evidence of difficulty. By way of example, one subject said FEP made him think of "vapor," which seems indicative of auditory discrimination trouble or a distorted sense of rhyme. The latter seems most likely for two reasons. First, the subject had been able to correctly enunciate the trigram and to pick it out when the experimenter said it. Second, observations made in both of the writer's studies of paired associates learning (the present study and the study reported by Otto, 1961a, 1961b) indicate that poor readers are not reluctant to make seemingly unlikely assocations that appear to be based upon varying kinds of mispronunciations. (See Otto, 1961b, for a more comprehensive discussion).

Asking subjects for reasons for ease or difficulty in making associations revealed nothing pertinent to the present study. A reason commonly given and representative of the others is that the square figure looks like a box which ryhmes with GOX and is therefore easy to remember. Of course, in many instances the subject could verbalize no reason at all.

TABLE I

POOR READERS' MEAN ACQUISITION TRIALS: GRADE x REINFORCEMENT

Grade	Auditory	Visual	Kinesthetic
2	22.50	24.20	17.00
4	14.75	13.00	14.25
6	12.75	10.50	10.50

TABLE II

MEAN ACQUISITION TRIALS MODE OF REINFORCEMENT

A	V	K
10.4	8.7.	8.5

Summary

An earlier study by the writer showed that good, average, and poor readers differ significantly in ability to learn a list of five paired associates. The present study attempted to determine whether the performance of the poor readers was interfered with by 1) lack of ability to discriminate the stim-

ulus figures and the response trigrams and/or 2) an additional task of learning to pronounce the printed trigrams.

On the basis of the present data, both notions were rejected. It was concluded that an explanation for the findings of the earlier study must be sought elsewhere. Some leads for further study were pointed out.

REFERENCES

1. Edwards, A. L. Experimental Design in Psychological Research. New York: Rinehart, 1950.

2. Otto, W. "The Acquisition and Retention of Paired Associates by Good, Average, and Poor Readers," Journal of Educational Psychology, 1961a, 52, 241-248.

3. Otto, W. The Acquisition and Retention of Paired Associates by Good, Average, and Poor Readers", unpublished doctoral dissertation, University of Wisconsin, 1961b.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 8, April 1963)

Readers Who Lacked Word Analysis Skills: A Group Diagnosis

MILDRED C. ROBECK
University of California

THE ELEMENTARY school child who lacks skill in word analysis is seen by his parents as one who cannot "sound out" unfamiliar words. Teachers become perplexed when they observe his inability to remember sound-letter associations and to apply the principles of word structure which have been learned by other children of comparable age and ability.

Problem

The purpose of this study was to discover some of the pertinent characteristics of reading clinic children whose most apparent difficulty was the lack of skill in word attack.

Problem 1: To analyze and describe the oral reading characteristics and error patterns of selected reading-clinic children.

Problem 2: To test the null hypothesis that readers who lack word analysis skill show no significant differences in intellectual functioning from a) the Wechsler population sample, b) the reading clinic population.

Problem 3: To compare reading achievement and reading potential as these are indicated by standardized aptitude, capacity, and achievement tests.

Problem 4: To isolate clues revealed in the backgrounds of the children which might be explored further as causal factors.

Subjects

Twenty reading-clinic cases whose major reading difficulty was a lack of word attack skill were identified. Subjects were 19 boys and one girl who ranged from six-years, eleven months to twelve-years, one month in age, second to seventh grade, and 98 to 136 IQ on the Wechsler Intelligence Scale for Children (WISC). Excluded from the study were children with extreme physical handicaps, limited general intelligence, bi-lingual home environment, and apparent emotional involvement.

Procedure

Oral reading samples, 77 to 266 words in length, were recorded on magnetic tape. Errors were transferred symbolically to oral reading sheets. Tabulations were made of each of the following errors: insertions, stops, repetitions, substitutions, refusals, omissions, and reversals. Allowance for differences in the length and difficulty of the samples was made by converting numbers of errors to percentages. Means and standard deviations were determined for six kinds of errors. Samples from children who deviated less than one SD in any one kind of error were analyzed for typical error patterns.

WISC's were administered to 80 Reading Clinic children enrolled over a two-year period. Subtest deviations were calculated from each subject's own scaled score mean. This had the effect of partialing out the general intelligence factor (1). Sigma scores were based on the $\underline{SE}_m$ appropriate to the child's age at the time of the test (3). The t's were determined for the experimental group of 20 cases and for the remaining 60 cases in the clinic population.

Children in the primary section of the reading clinic were administered Monroe's Reading Aptitude Tests, Primary Form, and Gates Advanced Primary Reading Tests, Type 1 and Type 2. Children in the upper elementary sections were administered Durrell-Sullivan Reading Capacity and Achievement Tests, Form A or B. Error ratios in oral reading were compared with silent reading achievement and with reading capacity.

Contributing factors

Cumulative folders for each of the 20 children in the sample were examined for factors related to reading difficulty. Files included pertinent information from school records, notes on interviews with parents, some follow-up tests which were not given to all subjects, and observations of the child's reactions to the clinic program. Factors were tabulated and arranged in order of frequency.

Results

Twenty children whose major reading difficulty was lack of word analysis skill made few or no insertions, omissions, and reversals. Fluency errors represented about half of the total, with stops twice as frequent as repetitions. Almost half of the errors were on word identification. The most frequent of these were substitutions (28 percent) and refusals (18 percent).

Nine of the children deviated less than one sigma in each error category. Their error patterns were regarded as "typical". They tended to phrase well at times; but lapsed into some word-by-word reading when recognition errors accumulated. The substitutions usually were words with like beginnings, similar configuration, and made sense in the preceding context. There was little evidence of inspection of the remainder of a sentence for context clues. Refusals occurred when no word analysis clues were available or when meaning had been destroyed by previous errors. Half of the children frequently asked for verification of words.

Repetitions occurred when reading did not make sense or when context clues were needed. Subjects stopped both before and after substitutions frequently. Some children tended to produce the first sound or syllable and then stopped until the word was supplied.

Intellectual Functioning

The experimental group and the reading-clinic controls showed significant strength, beyond the .001 level, in Comprehension, Similarities, Vocabulary, Picture Completion, and Block Design (Table I). Compared with Wechsler's population sample, both groups showed significant weakness, beyond the .001 level, in Information, Arithmetic, Digit Span, and Coding. In another study, recently completed, this investigator found that problem readers tended to deviate in large numbers, toward both ends of the continuum, in Picture Arrangement and Object Assembly subtests (2).

Reading Potential and Reading Achievement

The four cases assigned to the primary sections of the reading clinic averaged above the 50th percentile in each area measured by Monroe's Reading Aptitude Tests (Table II). In clinic practice, the general intelligence factor was assumed to operate in all tests; therefore, subtests which indicated aptitude at 25 or more percentile points below general intelligence were considered clues for further observation. By this criterion two children were low in visual aptitude, three were low in auditory aptitude, and one was low in articulation.

The same children averaged one-half grade below their school placement in silent reading achievement (Table III). Oral reading analysis, however, indicated an average recognition error for every six running words on material which was a full grade below school grade placement. This ratio was far below the accepted criterion of 1:20 at the instructional level.

Sixteen children, whose M grade placement was 4.7, averaged 6.4, on the Durrell-Sullivan Reading Capacity Test. They averaged more than one and one-half years above their M chronological age of ten years, two months. This result is not surprising since the clinic sample rated significantly high on the WISC Vocabulary and Comprehension subtests, after above average intelligence (108.5) had been equated.

On the Durrell-Sullivan Reading Achievement Test, clinic subjects averaged 2.5, more than two grades below school placement and 3.9 grades below indicated capacity. Their indicated reading potential averaged four years above their reading age of seven years, nine months.

On oral reading materials, which were comparable in difficulty to the indicated silent reading achievement level, they averaged one word recognition error for each six running words.

Possible Causal Factors

The review of Case Histories revealed combinations of factors which may have contributed to the reading difficulty. Of 93 factors recorded, those which appeared in frequencies greater than three were the following: Minimum entrance age- twelve; Poor auditory memory- eleven; Poor visual memory- ten; Lack of auditory discrimination- eight; Mixed lateral dominance- eight; Social immaturity- six; Delayed articulation- four; Distractability - four.

Summary

1. The oral reading of children who lacked word attack skills showed a high proportion of fluency errors but was not characterized by undue tension associated with the reading situation. The most frequent errors of recognition were substitutions of words with like beginnings and similar configuration which made sense with preceding context.

2. The null hypothesis that readers who lacked word-analysis skill were not significantly different in intellectual functioning from the Wechsler population sample was rejected for all subtests except Object Assembly. Selected clinic cases were significantly high (.001) in Comprehension, Similarities, Vocabulary, Picture Completion, Block Design, and in Picture Arrangement (.01 level). They were significantly low (.001) in Information, Arithmetic, Digit Span, and Coding.

The null hypothesis that readers who lacked word analysis skill were not significantly different in intellectual functioning from other problem readers held for all subtests except Picture Arrangement and Object Assembly.

3. The experimental group indicated reading potential beyond age-grade expectancies as measured by standardized aptitude and reading capacity tests. Primary age children indicated one-half grade retardation and upper grade children showed more than two grades retardation in silent reading.

achievement tests. All subjects produced error ratios at the frustration level when reading orally from graded materials comparable in difficulty to individual silent reading achievement levels.

4. Of the factors suspected as contributing causes, three appeared in frequencies approximating half the cases investigated: minimum entrance age, poor auditory memory, and poor visual memory.

REFERENCES

1. Cohen, Jacob. "The Factorial Structure of the WISC at Ages 7-6, 10-6, and 13-6," Journal of Consulting Psychology, XXIII (1959), pp. 285-99.
2. Wechsler, David. Wechsler Intelligence Scale for Children, Manual (New York: The Psychological Corporation, 1949).
3. Robeck, Mildred C. Intellectual Strengths and Weaknesses Shown by Reading Clinic Children on the WISC, Report presented to the Workshop in Reading Research, Chicago, April 17, 1962.

TABLE I

WISC SUBTEST DISTRIBUTIONS OF t

Subtest	Experimental Group*		Clinic Population**	
	t	Probability	t	Probability
Information	-11.26	.001	-55.98	.001
Comprehension	35.44	.001	85.68	.001
Arithmetic	-29.79	.001	-99.32	.001
Similarities	24.92	.001	81.95	.001
Vocabulary	19.58	.001	69.12	.001
(Digit Span)	-28.49	.001	-75.54	.001
Picture Completion	9.76	.001	43.32	.001
Picture Arrangement	3.62	.01	1.19	-
Block Design	6.30	.001	13.60	.001
Object Assembly	- 1.75	-	4.80	.001
Coding	-23.52	.001	-39.33	.001

* N = 20. Probability at .01 = 2.845; at .001 = 3.85.
** N = 60. Probability at .01 = 2.660; at .001 = 3.46.

TABLE II

RESULTS OF MONROE'S READING APTITUDE TESTS IN PERCENTILE RANKS

Name	Intelligence*	Visual	Auditory	Motor	Articulation	Language	M	R-L**
Jeff	98	57+	73+	95	82	78	73	17-0
Jane	95	64+	90	96	30+	95	75	15-2
Don	89	85	25+	72	89	99	74	17-0
Clint	70	60	16+	83	86	94	68	10-7
M =	88	67	51	87	72	92	73	

* WISC, converted to PR
** Hand-eye-foot preference
+ 25 or more PR points below IQ

TABLE III

ACHIEVEMENT LEVELS IN SILENT AND ORAL READING

Subjects		Gates Advanced Primary		Oral Reading	
Name	Grade	Word	Paragraph	Level	Error Ratio*
Jeff	3.0	2.2	2.1	1.8	1:4
Jane	3.0	2.5	2.0	1.8	1:6
Don	3.0	3.4	3.1	2.6	1:8
Clint	3.0	2.1	1.9	1.8	1:5
M =	3.0	2.6	2.3	2.0	1:6

*Recognition error: running words

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 8, April 1963)

How Intermediate Children Rank the Subjects

WAYNE L. HERMAN, Jr.
Laurel, Maryland

A LITTLE over a decade ago, Jersild conducted an extensive study of children's interests (1). A facet of his study entailed the interest children had in the subject areas. An open-end questionnaire was used in Jersild's study. Social studies generally ranked first as the subject children liked the least. The author was stimulated to investigate children's interests in the subject areas when a professor of a graduate course queried if social studies would rank low if an interest inventory were developed listing activities in the subject areas.

The Design of the Study

A multiple-choice Interest Inventory was developed using five subject areas: arithmetic, English, science, social studies and spelling. Each multiple-choice grouping had three items. A "Yes" was placed after the item which was liked the most and a "No" was placed after the item which was disliked the most. Forty groupings composed the Inventory. This is an example of one of the groupings.

(Yes or No)

Give a talk on the United States government______

Give a talk on insects ______

Give a talk on how multiplication tables help us______

For a basis of comparison, five open-end questions similar to Jersild's questions composed the last part of the questionnaire. Also, the children were asked to rank the subjects — arithmetic, English, science, social studies and spelling — in the order in which they liked them.

In November 1960, the test was administered by the author, a teacher in the school being tested, and was read in its entirety to preclude the reading problem of some pupils. The number participating in the study was 107 boys and 107 girls of the Caln Consolidated School in Thorndale, southeastern Pennsylvania. Three classrooms on each of the fourth-, fifth- and sixth-grade levels, or a total of nine classrooms, were administered the questionnaire separately. All of the subjects used in the study are taught in each heterogeneously grouped classroom. The grade distribution of the test population follows:

	Girls	Boys
Fourth grade	31	24
Fifth grade	42	38
Sixth grade	34	45
Totals	107	107

The Findings

When all the boys and girls in the test population are considered, social studies items in the Interest Inventory were rejected more frequently than were the items of any other subject. Social studies items were disliked more than any other subject items by fourth-grade girls and boys and fifth-grade boys. Science and social studies items were tied for being disliked the most by sixth-grade girls. The fifth-grade girls disliked science items most and social studies items next. The subjects liked the least are listed in percent according to the number of times chosen in the multiple-choice Interest Inventory.

107 Boys	English	.237
	Social Studies	.220
	Arithmetic	.192
	Science	.186
	Spelling	.165
		1.000

107 Girls	Social Studies	.264
	Science	.263
	Arithmetic	.193
	English	.170
	Spelling	.110
		1.000

214 Boys and Girls	Social Studies	.242
	Science	.224
	English	.203
	Arithmetic	.193
	Spelling	.138
		1.000

One of the more obvious findings was the interest strength which spelling has for most of the children. Girls gave spelling a wide lead over other subjects. Spelling, English and arithmetic in that order were the choices of most of the girls. Science, spelling and arithmetic in that order were top favorites for the boys. If the total test population is considered, spelling, arithmetic and English in that order are the favorite subjects. The subjects liked the most are listed in percent according to the number of times chosen in the multiple-choice Interest Inventory.

107 Boys	Science	.223
	Spelling	.219
	Arithmetic	.208
	Social Studies	.197
	English	.153
		1.000

107 Girls	Spelling	.304
	English	.231
	Arithmetic	.197
	Social Studies	.141
	Science	.127
		1.000

214 Boys and Girls	Spelling	.262
	Arithmetic	.202
	English	.192
	Science	.175
	Social Studies	.169
		1.000

When the children were asked which subject they liked the least, social studies and English were mentioned most frequently (number of responses are in parentheses): English (82), social studies (71), science (62), arithmetic (31), and spelling (19). When the boys and girls listed what they liked best in school, the number of responses were: spelling (137), arithmetic (89), social studies (54), science (42), and English (35).

The test population chose social studies fourth or fifth more times than any other subject. English drew a significant number of fourth and fifth place votes by the boys in the study. Girls placed science close to social studies as the subject they liked the least. Spelling was ranked highest by the girls followed by arithmetic and English. The boys ranked arithmetic first, spelling second and science third.

Conclusions and Recommendations

The results of this study compare favorably with Jersild's results. Admittedly, the validity of the Interest Inventory results depends on whether balance was achieved in the grouped items. Children have little interest in the subject of social studies; girls have less interest than boys. It has been said that social studies is the most monotonous of the subjects taught in the elementary school. Can the blame for this indictment be laid on the teacher? The Social Studies Council of Allegheny County in Pennsylvania has compiled a booklet of over a hundred activities which its teachers are using to inject enthusiasm, variety and vigor into its social studies' program.

The social studies book is usually voluminous. The reading problem for many of the children is not unreal. Study guides, rewriting the text down, taping reading assignments, grouping, the "buddy system," and texts on different reading levels are ways of dealing with the reading problem. The archaic mode of having each pupil take his turn reading orally should be eliminated.

The "expanding environment" concept of teaching social studies could be responsible for the low interest. Preston thinks the "expanding environment" concept is inconsistent with the development of children's social concepts — a development which is not orderly. Others suggest the children already know "expanding environment" content because of contemporary communication media — in the main, television. The home-school-neighborhood-city-state-nation-world plan is totally geographic. Preston suggests these sources of content: Community, Social Processes, Regions and Cultures, and the Past (2).

Although it is unrealistic to think we can produce high interest in all areas by most of the children, it is desirable that high interests of children be utilized to their optimum advantage; also, that an attempt is made to raise low interest levels by the introduction of a variety of teaching modes and activities, and by the evaluation of current theories of social studies scope and sequence.

REFERENCES

1. Jersild, Arthur T. Children's Interests and What They Suggest for Education (New York: Bureau of Publications, Teachers College, Columbia University, 1949).
2. Preston, Ralph C. Teaching Social Studies in the Elementary School, revised edition (New York: Rinehart and Company, Inc., 1958).

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 8, April 1963)

An Investigation of the Attitudes of Parents of Unsuccessful and Successful Readers*

DOROTHY P. MACDONALD**
Veterans' Administration
Dayton, Ohio

THE READING PROBLEM is universal and the research in the field is overwhelming, yet research leaves many questions unanswered concerning the abilities which influence the child's learning to read. Failure in reading has frequently been associated with emotional problems and, in turn, emotional problems have been attributed to reading difficulties. Disagreement persists on the question of whether the emotional problems are a cause or a result of the reading disability. It seems that an effective approach to the reading problem should involve the total personality of the child. It is doubtful that any conscientious investigator of the reading problem today would be naive to the extent that he would posit one specific cause for reading deficiency. With increasing emphasis on the "whole" individual, the atomistic approach is giving way to the more meaningful global viewpoint. Within this frame of reference, the child's relation with his parents certainly plays an important part in the development of his personality and thus his reading.

While counseling parents of URs at the Psychological Service for Children at the University of Denver, the writer observed that many of the parents were emotionally disturbed. She began to wonder what part, if any, did the attitudes toward their children play in the etiology of the UR. Very little research has been done in the area of parental attitudes regarding ways in which children should be treated. It is with this one area that this research is concerned. It is the purpose of this investigation to determine whether the attitudes toward child-rearing practices of the parents of URs differ significantly from the attitudes of parents of successful readers (SRs).

The Ss in this study were parents from three different groups of children. One group consisted of parents of URs at the University Children's Clinic. The other two groups were parents of URs and SRs in a public school. All groups were fairly well matched for age, education and socio-economic level. The children's groups were matched for IQ and reading ability.

For the collection of data, permission was granted by Bell and Schaefer (2) to use their Parental Attitude Research Instrument (PARI). With the exception of the parents of one UR, all parents in the clinic group were seen and given the PARI at the clinic. The writer administered the PARI in each home of the children in the two control groups. All children were given the WISC and Wide Range.

A chi-square test was computed as a means of testing the significance of difference of parental agreement between the three reading groups. Although the findings of this study did not support the hypothesis, it is felt that there were interesting implications.

The parents of the URs in the public school expressed attitudes that were more similar to those expressed by the parents of SRs than by parents of the URs in the child guidance clinic. Significant differences in attitudes between the mothers of the three groups were found in four of the PARI scales: Breaking the Will, Strictness, Suppression of Aggression, and Suppression of Sex. The differences were found mainly between the mothers in the clinic and the mothers of poor readers in the public schools. The clinic mothers were inclined to be less strict, less likely to suppress aggressive acts of their children, and less likely to suppress their children's curiosity about sex than the mothers of the URs and SRs in the public school.

Significant differences in attitudes between the fathers of the three groups were found in three of the PARI scales: Excluding Outside Influences, Suppression of Aggression, and Suppression of Sex. The differences were found mainly between the fathers in the clinic and the fathers of URs in the public school. The clinic fathers were inclined to be less likely than the public school UR or SR fathers to suppress their children's curiosity about sex, and less likely to feel that children should never learn things outside the home which would make them doubt their parents' ideas.

---

* This paper is a condensation of a dissertation submitted in partial fulfillment of the requirements for the Ph.D. degree at the University of Denver, Denver, Colorado.

** Now at Bethesda Hospital, Denver, Colorado, as clinical psychologist.

The means of the scores of the clinic fathers tended to be predominantly the lowest of all parents. Owing to the fact that the URs in this study were all male, it may prove fruitful if future UR studies examine the relationship between the UR and his father. Very little research has been devoted to the attitudes and personality traits of fathers.

BIBLIOGRAPHY

1. Axline, Virginia. "Nondirective therapy for Poor Readers," Journal of Consulting Psychology, XI (1947), pp. 61-69.
2. Bell, Richard, and Schaefer, Earl. The Parental Attitude Research Instrument. National Institute of Mental Health, PHS, U.S. Department of Health, Education and Welfare (1957).
3. Blanchard, Phyllis. "The Psychogenic Factors in Some Cases of Reading Disability," American Journal of Orthopsychiatry, V (1935), pp. 361-375.
4. Blanchard, Phyllis. "Reading Disabilities in Relation to Difficulties of Personality and Emotional Development," Mental Hygiene, XX (1936), pp. 384-413.
5. Graham, E. Ellis. An Exploration of a Theory of the Emotional Bases for Reading Failure, unpublished doctorate dissertation, University of Denver (May 1951).
6. Johnson, Marjorie S. "A Study of Diagnostic and Remedial Procedure in a Reading Clinic Laboratory School," Journal of Educational Research, XLVIII (1955), pp. 556-578.
7. Lehman, Edwin. An Exploratory Study of Maternal Overprotection in Relation to Reading Retardation, unpublished master's thesis, University of Denver (1957).
8. Levy, D.M. "Relation of Maternal Overprotection to School Grades and Intelligence Tests," American Journal of Orthopsychiatry, III (1933), pp. 26-34.
9. Missildine, W.H. "The Emotional Background of Thirty Children with Reading Disabilities with Emphasis on Its Coercive Elements," The Nervous Child, V (1946), pp. 263-272.
10. Monroe, Marion. Children Who Cannot Read (Chicago: University of Chicago Press, 1932).
11. Preston, Mary. "The Reaction of Parents to Reading Failure," Child Development, X (1939), pp. 173-179.
12. Schmitt, William. The Dynamics of Parent-Child Relationships, unpublished master's thesis, University of Denver (1952).
13. Shatter, Florence. "An Investigation of the Effectiveness in Group Therapy Program Including the Child and His Mother for Remediation of Reading Disability," Dissertation Abstracts, XVII, No. 5, University Microfilm (Ann Arbor, Michigan, 1957), p. 1032.
14. Stewart, Robert. "Personality Maladjustment and Reading Achievement," American Journal of Orthopsychiatry, XX (1950), pp. 410-417.
15. Sutton, Rachel. "A Study of Certain Factors Associated with Reading Readiness in the Kindergarten," Journal of Educational Research, XLVIII (1955), pp. 513-538.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 8, April 1963)

Age, I Q, and Improvement of Reading*

NORMAN M. CHANSKY

North Carolina State College, Raleigh

CHILDREN with high IQ's, it has been assumed by many reading specialists, make greater progress in a Remedial Reading program than children with lower IQ's. The belief stems, perhaps, from the high positive correlation between achievement and intelligence as currently measured. The potential to profit from remedial reading is held to be greater for the children with the higher IQ's. The resulting practice has been to select for Remedial Reading children with the higher IQ's.

In a previous study, Chanksy (1) found an inverse relationship between scholastic aptitude and improvement in reading comprehension and in vocabulary. The subjects were college students. The purpose of the present research was to study the effect of mental ability on the improvement of reading in elementary-school children. The subjects were 41 children from the Highland, New York, Central School, registered in the Remedial Reading class for the academic year. The subjects met in small groups four times weekly for a period of 30 minutes. Alternate forms of the Gates Reading Survey were administered in pre- and post-sessions. Growth was reported in months. The California Short Form Test of Mental Maturity was administered to the children at the beginning of the school year. This test yields ratio IQ's.

The age range of the students was eight to 14 years, with a median age of eleven. The IQ range was 82 to 118 with a median IQ of 100. The range of changes in reading scores was -2 to +33 months, with a median change of 13 months.

Ages, IQ's, and reading-change scores were ranked and tau coefficients were computed. The correlation between IQ and reading improvement was +.17, between IQ and age -.49, and between age and reading improvement -.32. The latter correlations are those expected, younger children responding to instruction better than older ones. Younger children had, perhaps, substantially higher IQ's. When the effect of age was partialled out of the interrelationship between IQ and reading improvement, the result was a tau coefficient of +.01. Age, or an unstudied correlate of it, was responsible for the very slight relationship between intelligence and improvement. When IQ was partialled out of the correlation between age and improvement, the result was a change to a tau coefficient of -.35.

In summary, no empirical support was found for the belief that children with high IQ's make the greatest progress in Remedial Reading. Rather age, or some correlate of it, perhaps history of failure or responsiveness to small-group instruction, was found to be related to improvement in reading. It appears to the writer that there is a need to seek a criterion other than mental ability in making selections for classes in Remedial Reading.

REFERENCE

1. Chanksy, N. M. "Improvement of Reading in College," Journal of Educational Research, LI (1957), pp. 313-317.

---

* The author wishes to thank Mr. Donald Baines, Elementary School Principal, Highland Central Schools, Highland, New York, for his assistance in collecting the data.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 8, April 1963)

The Relationship Between Pupil Self-Evaluation, Teacher Evaluation of the Pupil and Scholastic Achievement

IRA J. GORDON and PATRICIA C. WOOD*
University of Florida

SELF THEORISTS have proposed a fundamental relationship between self-concept and behavior. In addition, they have postulated that threat to the self decreases the ability of the person to perceive both self and situation accurately.

Previous research leaves unresolved the relationship between teachers' behavior and the self-concept of individual pupils. Although it is a widely held belief among educators and many psychologists that such a relationship exists, empirical data are needed. Perkins (3) and Staines (5) both found relationships between teacher behavior and self-conceptualization in children. Sears (4) found that there was a relationship between pupil self-ratings on mental ability and teacher ratings of mental ability on youngsters in middle childhood.

The research on accuracy of self-evaluation by children suggests first, that there is a wide range in ability to view self realistically (1), second, that there seems to be a tendency to overestimate (6) and third, that accuracy seems to be a function not only of the variable being estimated (6) but also the degree of "threat" (3). A fourth generalization from the research has been stated by Wylie(6): that there is some limited evidence of individual consistency of self-rating, but a lack of detailed information.

This study is designed to explore the nature of the agreements, or disagreements, between pupil and teacher estimates of the pupil and the relationships of these evaluations to scholastic achievement. What happens when teacher and pupil agree? Is the pupil then more able to evaluate his performance realistically? Is teacher-pupil disagreement related to more perceptual distortion on the part of the pupil? Theoretically, there should be a relationship between the teacher's and the pupil's estimate. Is this borne out empirically?

Further, are the pupils in this study individually consistent? Does the youngster who overestimates himself on a general self-rating scale on a variety of school items also overestimate himself on his estimate of performance on a specific task? Is there a general tendency, throughout the group, to overestimate?

Procedure

Three classes of fifth graders and two classes of sixth graders (72 boys and 76 girls) in a metropolitan Florida school and their five teachers served as subjects. These teachers averaged 15 years of experience, were all women, two held master's degrees, which the other three held bachelor's degrees in elementary education.

In the spring of 1960, the pupils rated themselves on the 15 school items of the "How I See Myself" 5-point scale developed by Gordon from the Jersild (2) categories. The following are a sample of the fifteen categories:

In school, I don't stay with things and finish them	1 2 3 4 5	In school, I stay with something till I finish
I don't do well in school work	1 2 3 4 5	I do well in school work
School isn't interesting to me	1 2 3 4 5	School is very interesting to me
I'm not as smart as as the others in my class	1 2 3 4 5	I'm smarter than most of the others in my class
I don't read well	1 2 3 4 5	I read very well

The children were given the following explanation of the questions:

"I would like for you to help me with a study I'm making of teacher-pupil relationships — the way you and your teacher work and study together. These sheets are for my use only. None of your classmates, or your teacher, or the principal, may see them - this is just between you and me.

"There are fifteen items on this sheet that I want

---

* The writers wish to acknowledge the help of Dr. William D. Spears on this project.

you to mark according to the way you feel most of the time. No one feels the same about school all the time, but try to think of the way you feel most of the time. Please remember that no one else will see your papers."

The children were given time to ask questions and then each item was read, and the child circled the number that represented his feeling on the item.

In order to avoid "set", about half the items, randomly distributed were presented with one as the positive answer, and five the negative. For scoring purposes, all positive items were ranked as five. A total score on the fifteen items was computed for each child. The children were listed from highest to lowest using this score, then this list was divided into stanines.

Each teacher was asked to rank his pupils from highest to lowest on the same self-report scale, then this list was divided into stanines. The two lists were compared. If the teacher placed the child in the same stanine as he placed himself, they were said to agree. If the teacher did not place the child in the same stanine as he placed himself, they were said to disagree.

Two weeks later, the children were asked to rate themselves as to their estimate of their relative position in class achievement, as measured by the Stanford Achievement Test. They were told that in a few weeks they would be given the regular yearly achievement tests. They were asked to consider themselves in relation to their classmates – "the number of people in the room you think are smarter than you, and then the number of people you think are not as smart as you are generally." They were asked to write on the paper given them the number that represented their class position, whether they thought they were first, or second, or fourth, or sixteenth, or thirty-fourth. These estimates were arranged in order from highest to lowest in each class then each list was divided into stanines.

The teachers were asked to rank their classes from highest to lowest on achievement as measured by the Stanford Achievement Test. The lists were divided into stanines and compared to the child's estimates. The same criterion as before was used for agree and disagree.

Two weeks later, each teacher gave the intermediate battery, Form JM, of the Stanford Achievement Test. The scores of the pupils were listed from highest to lowest and divided into stanines. If the teacher or the child placed the child in the same stanine on self-report scale or achievement estimate as his actual achievement test score, he (the teacher or the child) was said to be accurate. If he did not place the child in the same stanine as the achievement test score, he was said to be inaccurate. In cases where the teacher and the child disagreed, either one could be accurate by placing the child in the same stanine as his achievement test score.

For example, if the child fell in the third stanine on the self-report scale and the third stanine on his actual achievement test score, he was said to be accurate on his self-report. The same definition applied to test estimate.

The hypotheses to be tested were:

1. There will be a positive relationship between the child's agreement with the teacher on the self-report scale and his accuracy on the self-report scale as an estimate of achievement.
2. There will be a positive relationship between the child's agreement with the teacher on estimate of test ranking and his accuracy in ranking himself.
3. When the teacher and child are inaccurate, they will agree on the direction of inaccuracy (over or under-estimate) in comparing self-report scale to actual achievement test rank.
4. When the teacher and the child are inaccurate on the achievement estimate, they will agree on the direction of the inaccuracy.
5. When the teacher and the child disagree on the self-report scale, the teacher's ranking will be closer to actual ranking on achievement test.
6. Where the teacher and child disagree on the estimate of achievement, the teacher's ranking will be closer to actual rank on achievement test.
7. Children who err in their estimate on one measure will err in the same direction on the other measure.

Results

The data for all the hypotheses, except the fifth and sixth, were analyzed by means of the chi-square test of significance.

Table I presents the findings on the teacher-pupil hypotheses in which chi-square was used. The first hypothesis, that there will be a positive correlation between agreement on self-report and the child's accuracy on ranking himself in the same relative class position on the self-report scale that he holds on achievement test performance, was upheld. The number who were accurate was small, but of those who were, twice as many also had teacher-pupil agreement.

High accuracy was not expected for two reasons. First, agreement on a stanine classification is a rigorous test. In a class of 27 pupils, the stanine distribution is: 1, 2, 4, 5, 7, 5, 4, 2, 1. The child could be ranked second on self-report and actually be third in achievement and be classified as "inaccurate." Second, the child was rating himself generally on many school items, including attitudes toward school subjects and teachers as well as ability. This rating was being used against test achievement on a standardized test.

When the children were asked to rank themselves on the particular task, the number who were inaccurate doubled but was still less than one-third of the group. There was no relationship between agreement and accuracy on the specific task.

Since teacher and pupil both tend to be inaccurate by this rigorous definition, is there any pattern to their inaccuracy? Three types of hypotheses were explored, one dealing with direction of error, the second dealing with amount, and the third with consistency of error. Hypotheses three and four (direction of error) were upheld (see Table I). We cannot assign causation, but the high degree of relationship is noteworthy.

Hypotheses five and six (amount of error) were tested by use of the *t* test for correlated means. The mean difference between the discrepancy scores (actual stanine - estimated stanine) for both pupils and teachers on both self-report and test estimate were computed. In both cases, the teachers were closer to the actual rank than were the pupils. The differences between teacher and pupils were significant at the .001 level. That is, even though the teachers missed the correct rank, they were closer to being "on target" than were the pupils.

The seventh hypothesis, that pupils would be consistent in their over- or underestimation was upheld (see Table II). The child who overestimated his achievement test score also tended to overestimate himself on the self-report scale.

The data were analyzed for sex differences, but none were found. Teachers were neither more nor less accurate on boys than girls; neither more nor less in agreement with one sex or the other. Neither boys nor girls had higher accuracy, nor different tendencies to over- or underestimate.

Discussion

In this study, teachers were not able to predict with much success the way children would evaluate themselves on self-reports about school and estimate of test performance. There was agreement on direction of error of evaluation, but little on actual ranking. It may be that the teacher's evaluation of the child influences the way he evaluates himself. We seem to have an "over-reaction" situation. The teacher overestimates the child a little, the child overestimates himself even more. Although we cannot state a cause-effect relationship, Staines' (5) work would suggest this possibility.

The relationship between agreement and accuracy on the self-report tends to support the idea that agreement might be threat-reducing leading the child to see himself realistically. Disagreement may lead to the tendency for the pupil to distort, hence the over-reaction.

Only seven children were accurate on both estimates. Three of these had complete agreement with the teacher, and three more agreed with the teacher on one estimate. Of the eight others who were accurate on self-report but either over or under-evaluated their test estimate, none had complete agreement with the teacher, three had no agreement, and only two agreed with the teacher's estimate of achievement. Although these numbers are too small for any statistical meaning, they do reflect a possible trend of a positive relationship between pupil-teacher agreement and pupil accuracy of self-estimate.

We found this group of pupils to be consistent in both under- and over-estimation. We found no more tendency to overestimate than to underestimate. Further, if we couple this pupil consistency with the significant relationship to teacher estimate, there is a pattern of consistent error by both teacher and child. Although the magnitude of the teacher's error is significantly less than the pupils', the consistency of direction seems most significant. The big question, of course, is who is influencing whom? Is it that the teacher understands how the child sees himself, or does the teacher's evaluation shape the child? We cannot say from this study. We can say that the child's general view of self in relation to school is reflected in his specific view of his ability on a standardized achievement test.

Summary

Teachers and children in five fifth-and sixth-grade classes were asked to estimate the child's performance on a standardized achievement test. The children also rated themselves on a "How I See Myself" scale on general feelings about themselves in relation to the school situation. The teachers estimated the rank order of the children on these reports. All ranks were placed in stanines, and the amount of agreement between teacher and child, and degree and amount of agreement between self-reports, estimates, and actual test rank were analyzed. Relationships were found between pupil-teacher agreement on self-reports and accuracy of the self-report as an estimate of actual test rank and between direction of error on both self-report and test estimate between teacher and pupils. The pupils were consistent in the direction (over- or underestimation) on both measures. Teachers were closer to correct estimates of test ranks than were pupils. Teacher-pupil disagreement seems to be related to perceptual distortion in the same direction by both. The role of threat was discussed as a possible factor in the amount and degree of pupil error.

REFERENCES

1. Brandt, Richard M. "The Accuracy of Self Estimate: A Measure of Self-Concept Reality," *Genetic Psychology Monographs*, LVIII (1958), pp 55-99.
2. Jersild, A. T. *In Search of Self* (New York: Bureau of Publications, Teachers College,

Columbia University, 1952).

3. Perkins, Hugh V. "Factors Influencing Change in Children's Self Concepts," *Child Development* XXIX (1958), pp. 221-30.

4. Sears, Pauline. "The Pursuit of Self-Esteem: The Middle Childhood Years," *Newsletter*, Division of Developmental Psychology (Fall, 1960), pp. 1-5.

5. Staines, J. "The Self-Picture as a Factor in the Classroom," *British Journal of Educational Psychology* XXVIII (1958), pp. 97-111.

6. Wylie, R. *The Self-Concept* (Lincoln: University of Nebraska Press, 1961).

TABLE I

TEACHER AND CHILD ESTIMATES AND SCHOLASTIC ACHIEVEMENT RATING

Variable	Number of Children		
1. Agreement between Teacher and Pupil on Self-report	Accurate	Inaccurate	Chi-Square
Agree	10	26	13.46*
Disagree	5	105	
2. Agreement between Teacher and Pupil on Achievement Estimates			
Agree	13	29	1.38
Disagree	21	83	
3. Direction of error of estimate, Self-report	Overestimate	Underestimate	
Teacher overestimates	25	14	19.08*
Teaches underestimates	10	32	
4. Direction of error of estimate, achievement			
Teacher overestimates	25	12	16.25*
Teacher underestimates	10	25	

* Difference is significant at the .01 level of confidence

TABLE II

CONSISTENCY OF CHILDREN'S ESTIMATE BETWEEN SELF-REPORT AND TEST ESTIMATE

Direction of error of estimate on rank achievement test	Direction of error, self-report, to actual rank on achievement		Chi-Square
	Overestimate	Underestimate	
Overestimate	40	11	34.6**
Underestimate	11	42	

**Difference is significant at the .001 level of confidence

Field News

SOME FINDINGS OF PROJECT ENGLISH AND PROJECT TALENT

Some of the findings to date are: High school boys have larger vocabularies than girls, but girls have better reading comprehension and effective expression: more than one-fourth of the high school freshmen outrank the average high school senior in English; high English scores almost always are accompanied by high mathematic scores; the average high school senior can answer correctly only 28 percent of the questions based on the level of difficulty found in *Atlantic Monthly* or *Saturday Review*, he does better (44%) with *Readers Digest* and gets up to 54 percent correct with *Saturday Evening Post*.

PARENTS FAVOR NATION-WIDE TEST

The American Institute of Public Opinion (Gallup Poll) reports that 77 percent of parents would favor a standard nation-wide test to be given to all students when they enter and when they graduate from high school. The question was phrased on the assumption that such a test would measure achievement and would permit comparison among graduates of various high schools.

PHYSICAL FITNESS AND ACADEMIC SUPERIORITY

From a survey undertaken in the public schools of Springfield, Missouri, the President's Council on Youth Fitness concludes that students who maintain a high level of physical fitness tended, on the average, to have the best grades, the fewest absences, the heaviest load of extracurricular activities, and the most healthy attitudes as judged by their teachers.

CREATIVITY - INHERITED OR ACQUIRED?

Sidney J. Parnes of the University of Buffalo reports in *Teachers College Record*, January 1963, that the hypothesis is supported that creativity can be learned. In a carefully matched experimental study, it was found that creative problem-solving students showed substantial gains on two tests of idea-*quantity*; they showed clear superiority over the control group in three tests of *qualitative* production of ideas; and they were characterized by the personality characteristic, dominance.

PAUL R. MORT - RESEARCH FELLOWSHIPS

Twelve research fellowships in the Institute of Administrative Research, Teachers College, Columbia University, are announced for the academic year 1963-64. Each fellowship carries a stipend of $2,440 for the academic year, including tuition exemptions for eight semester hours each semester. Further information may be obtained from the Institute of Administrative Research, 525 West 120th Street, New York 27, New York.

TERMAN'S GIFTED CHILDREN

Dr. Helen Marshall, psychologist at Stanford University, has completed a new study of 163 of Terman's 1500 gifted children. Some of her findings are: the income of the men rose from $2,675 to $13,500 in the 1940-60 years; that of the women $2,287 to $7,800 for the same period; politically, the Republicans outnumber the Democrats, two-to-one; and their greatest satisfactions are work, marriage, and children - in that order.

TEACHER EDUCATION AND PROGRAMED INSTRUCTION

The American Association of Colleges for Teacher Education is conducting a survey of current research and experimentation on programed learning in teacher education. About 50 percent of the AACTE member institutions have identified 600 faculty members working in one or more areas of programed learning. It is to these individuals that a questionnaire is being sent to determine the status of programed learning.

ACCELERATED COLLEGE EDUCATION

Nassau Community College, a two-year liberal arts college, has instituted an accelerated program somewhat similar to the trimester plan currently in operation at some colleges. At Nassau the program is known as the "Two Plus Two Plan." It consists of the usual fall and spring semesters plus a double summer session offering. The double summer session enables students to take as many as four semester length-courses or one or two complete full-year courses in order that the Associate Degree may be earned in less than two years.

UCEA RESEARCH SEMINAR

The University Council for Educational Administration announces that the University of Illinois will be host for a three-week seminar on "Community Decision Making and Change and Their Influence on Education" in late summer of this year. The seminar is an outgrowth and extension of the work initiated in 1961 by the UCEA Task Force on The School as a Political Institution. The seminar will be supported by a development grant from the Cooperative Research Branch.

PROJECT SOCIAL STUDIES

The U S. Office of Education announces the initiation of Project Social Studies. This program is designed to improve research, instruction, teacher education, and the dissemination of information in the social studies. The Cooperative Research Branch is cooperating with other divisions within the office by making funds available to colleges, universities, state education departments, and through the latter, local school systems for (a) basic and applied research projects, (b) curriculum study centers and (c) research development activities in the social studies. These activities may be conducted at any level, from pre-school through adult education, and may focus on one or more of the social studies.

Although the areas referred to above are the ones in which funds are currently available, they represent only the initial efforts in the social studies. If you or any of your colleagues have ideas of other activities that might improve research, instruction, teacher education, or the dissemination of information in this field, the office invites you to express your interests. Further information and instructions for submitting applications may be obtained from Special Projects Area, Cooperative Research Branch, U. S. Office of Education, Department of Health, Education, and Welfare, Washington 25, D. C.

DROPOUTS AND JOBLESS RATE

The National Consumer Finance Association estimates that the unemployment rate among youth who left high school last year without graduating is 26.8 percent as compared to the 17. 9 percent rate of high school graduates. A study of young people whose high school class graduated in 1959 showed that dropouts had a jobless rate twice as high as that of high school graduates. Current Census Bureau figures show median incomes of men with college degrees as $7260; with high school diplomas, $5050; with grade school education, $3450. Dropouts are suffering a growing handicap.

COLLEGE EDUCATION AND STUDENT FINANCES

Two recently completed studies support what previous studies have repeatedly found- financing is a major factor in starting and completing college education.

In one of the studies, "The Duration of Formal Education for High Ability Youth," conducted by the National Science Foundation, it was found that among the upper 30 percent of 17-year olds, "the largest single reason for failure to enter college appears to be inadequate financial resources." Furthermore, lack of money caused up to one-half of the male dropouts and one-third of the female dropouts.

The other study was a survey of the 11, 000 winners of the Sears-Roebuck Foundation agricultural scholarships for freshmen in the years 1936-56. It showed that 78.5 percent of the recipients earned degrees at the bachelor's level and over. Also, it showed that more than 75 percent of those responding to the questionnaire felt that financial aid received was substantial or vital.

KENNEDY FOUNDATION VISITING PROFESSORSHIPS

Mr. and Mrs. R. Sargent Shriver and President Felix C. Robb have announced the establishment of Kennedy Foundation Visiting Professorships at George Peabody College for Teachers. The professorships will be awarded to outstanding investigators from fields that can contribute to a better understanding of mental retardation, such as genetics, bio-chemistry, physiology, psychology, education, pediatrics, psychiatry, sociology, and anthropology. Preference will be given to scientists who have not previously worked on mental retardation. The professor will be expected to give a graduate seminar on a topic of his choice and to study mental retardation with a view toward assessing current knowledge and defining research problems and strategies. Several appointments will be made each year on an international basis. Applications and nominations are being received for the academic years 1963-64 and 1964-65. For further information write Nicholas Hobbs, George Peabody College for Teachers, Nashville 5, Tennessee.

PAN AMERICAN UNION RESEARCH

The educational research unit in the Division of Education of the Pan American Union currently is conducting a questionnaire survey of the programs and institutions in Latin America offering correspondent studies. Further information may be obtained by writing to the Union, Washington 6, D. C.

Readers are invited to send items of research interest to John Schmid, Professor, College of Education, University of Arkansas, Fayetteville, Arkansas.

Among Those Present

NORMAN M. CHANSKY received his PhD from Columbia in 1958. Currently a supervising school psychologist in New York, he has taught at Adelphi and the Oswego College of Education.

WAYNE L. HERMAN, Jr., is head teacher at Westwood Elementary School, Coatesville, Pa. His BA is from Ursinus and his MEd from Temple, where he is currently completing course work for the doctorate.

(Mrs.) DOROTHY P. MACDONALD attended Duke, Wayne, Colorado College, and Chicago and has her PhD from Denver. Since 1961 she has been a clinical psychologist with the Veterans Administration at Dayton, Ohio.

Readers Write

I question the standards used in selecting some research reports. —Rupert N. Evans, associate dean, college of education, University of Illinois.

The JER seems to have some of the defects of some other learned publications. It is statistically overwhelming. I would humbly suggest a re-evaluation in this area, particularly with respect to the quality of creativity and originality of reports. —James T. Mehorter, assistant professor of educational psychology, University of Vermont.

Congratulations to you and the other people responsible for the "new look' for The JER. I believe it is making it far and above the top journal reporting educational research. —John H. Suehr, principal, Alamosa (Colorado) High School.

The JER is much improved recently, but I dislike the poor writing techniques of some authors and the unwieldy size of the magazine. All in all, it's a fine job and a major contribution to education. —Robert Allen, vice president, Famous Artists and Writers Schools, Westport, Conn.

What about misleading titles? For example, "What is TV Doing to Our Children?" This question is not answered in the article. —Maynard Bjorgo, periodicals librarian, St. Cloud State College, Minn.

I think The JER offers a good balance. While I do not have equal interest in all departments and aspects, nevertheless, it is good to be exposed to such material. Keep up the good work! —John W. McIntosh, San Joaquin County, Stockton, California.

I would appreciate more articles on the meaning of research for the classroom—Ruth C. Little, professional editor, department of school services, Wesleyan University.

MOVE, MOVING, MOVED? Won't you help us keep your Journal coming to you on time by keeping us posted on your whereabouts?

Your Name ________________________________

Old Address ________________________________

____

New Address ________________________________

____

DEMBAR EDUCATIONAL RESEARCH SERVICES, Inc.
Post Office Box 1148, Madison 1, Wis.

Executive Editors

WILSON B. THIEDE, Professor of Education, Educational Administration, and Agricultural and Extension Education, The University of Wisconsin, Madison, Wisconsin. Chairman.
THEODORE HARRIS, Professor of Educational Psychology and Education, The University of Wisconsin, Madison, Wisconsin.
L. JOSEPH LINS, Professor and Coordinator of Institutional Studies, The University of Wisconsin, Madison, Wisconsin.
PHILIP LAMBERT, Professor of Educational Psychology and Educational Administration, The University of Wisconsin, Madison, Wisconsin.
JOHN SCHMID, Professor of Education and Chairman of the Department of Educational Foundations, University of Arkansas, Fayetteville, Arkansas.

Editorial Board

JACOB O. BACH, Head, Educational Research Bureau, Southern Illinois University, Carbondale, Illinois.
WILLIAM A. BROWNELL, Dean, School of Education, University of California, Berkeley, California.
ROBERT A. DAVIS, Professor of Educational Research, George Peabody College for Teachers, Nashville, Tennessee.
GUY D. McGRATH, Dean, College of Education, Arizona State University, Tempe, Arizona.
HAROLD E. MITZEL, Director, Office of Research and Evaluation, Division of Teacher Education, College of the City of New York, New York City.
WILLARD C. OLSON, Dean, College of Education, University of Michigan, Ann Arbor, Michigan.
ARTHUR E. TRAXLER, Executive Director, Educational Records Bureau, New York 32, New York.

Associate Editors

HAROLD M. ANDERSON, Associate Professor of Education, University of Colorado, Boulder, Colorado.
VERNON E. ANDERSON, Dean, College of Education, University of Maryland, College Park, Maryland.
RUDYARD, K. BENT, Professor of Education, University of Arkansas, Fayetteville, Arkansas.
EMMETT ALBERT BETTS, Research Professor, University of Miami, Coral Gables, Florida.
LAWRENCE P. BLUM, Assistant Director, Student Personnel Services, University of Wisconsin-Milwaukee, Milwaukee 3, Wisconsin.
J. MURRAY LEE, Chairman, Department of Elementary Education, Southern Illinois University, Carbondale, Ill.
D. WELTY LEFEVER, Professor of Education, University of Southern California, Los Angeles, California.
S. E. TORSTEN LUND, Associate Professor of Education, University of California, Berkeley, California.
WILLIAM A. McCALL, Lecturer in Education, University of Miami, Coral Gables, Florida.
DELYTE W. MORRIS, President, Southern Illinois University, Carbondale, Illinois.
VALWORTH PLUMB, Chairman, Division of Education and Psychology, University of Minnesota (Duluth Branch), Duluth, Minnesota.
DAVID G. RYANS, Rand Corporation, Santa Monica, California.
LESTER B. SANDS, Professor of Education, University of California, Santa Barbara, Goleta, California.
MANDEL SHERMAN, President, and Consultant, Emotional Health Association, Los Angeles, California.
MILDRED B. STANTON, Consultant, Special Education, State Department of Education, Hartford 15, Conn.
W. W. THEISEN, Professor of Education, Marquette University, Miwaukee, Wisconsin.
M. R. TRABUE, Professor of Higher Education, University of Kentucky, Lexington, Kentucky.
J. WAYNE WRIGHTSTONE, Director, Bureau of Educational Research, Board of Education of City of New York, Brooklyn, New York.

Editors Emeriti

E. J. ASHBAUGH, Dean Emeritus, School of Education, Miami University, Oxford MR50, Ohio.
HARRY J. BAKER, 19050 Wiltshire, Lathrup Village, Michigan.
FOWLER D. BROOKS, Director Emeritus, Department of Psychology, DePauw University, Greencastle, Ind.
LEO J. BRUECKNER, Professor Emeritus, University of Minnesota, Minneapolis, Minnesota.
B. R. BUCKINGHAM, 33 Chatham Circle, Wellesley Hills, Massachusetts.
EDGAR A. DOLL, Director Emeritus, Devereux Schools, Devon, Pennsylvania.
HARL R. DOUGLASS, Director Emeritus, School of Education, University of Colorado, Boulder, Colorado.
E. DUNCAN GRIZZELL, Dean Emeritus, University of Pennsylvania, 640 Maxwelton Court, Lexington, Ky.
ERNEST HORN, Professor Emeritus of Education, State University of Iowa, W307 East Hall, Iowa City, Iowa.
EARL HUDELSON, Professor Emeritus of Education, West Virginia University, Morgantown, West Virginia.
A. M. JORDAN, Professor Emeritus, University of North Carolina, Chapel Hill, North Carolina.
ALFRED S. LEWERENZ, Director Emeritus of Evaluation and Research, 4107 Holly Knoll Drive, Los Angeles, California.
A. R. MEAD, Director Emeritus of Educational Research, University of Florida, 1719 N. W. 6th Ave., Gainesville, Florida.
M. J. NELSON, Dean Emeritus, Iowa State Teachers College, 6033 West Orange Drive, Glendale, Arizona.
JESSE B. SEARS, Professor Emeritus, Stanford University, Stanford, California.
CARLETON W. WASHBURNE, Chairman Emeritus, Department of Education, Brookl

The Journal of Educational Research

Dedicated to the Scientific Study of Education

VOLUME 56 MAY-JUNE 1963 NUMBER 9

It Says in the JER ---------- 450

Marks, Achievement, and Intelligence of Seventh Graders Who Were Retained (Nonpromoted) Once in Elementary School—*Constance K. Kamii and David P. Weikart* ---------- 452

The Discovery of Personality Traits—*John L. Horn* ---------- 460

Teacher Reactions to School Faculty Meetings—*Arthur Blumberg and Edmund Amidon* ---------- 466

A Comparison of the Achievement of Multi-Graded and Single-Graded Rural Elementary School Children—*Carmen J. Finley and Jack M. Thompson* ---------- 471

Personality Correlates of Attitude Toward Mathematics—*Lewis R. Aiken, Jr.* ---------- 476

An Evaluation of the California Achievement Test, Elementary Form W, Reading Vocabulary—*Charles Sheppard and William J. Campbell* ---------- 481

A Factorial Analysis of Fields of Scholarly Production—*Robert H. Knapp and Linda A. Ehlinger* ---------- 484

The Relation of Mental Health Analysis Scores to Teacher Ratings of Mental Health Status of Elementary Pupils—*Joseph C. Bledsoe* ---------- 488

The Effects of Chronological Age in Months on School Achievement—*Donald J. Dickinson and J. Donald Larson* 492

Intelligence and Gain in Reading as Related to Gains in the Sub-tests of the Stanford Achievement Test—*Carrie M. Scott* ---------- 494

Development of Responsibility Perceptions of Lower and Middle Class Children—*Michael Zunich* ---------- 497

The Use of Intelligence Tests—*Richard M. Sanders* ---------- 500

Field News ---------- 501

Among Those Present ---------- 503

The Journal of Educational Research is published at Madison, Wisconsin, monthly 10 times a year (the May-June and July-August issues being combined). The subscription price is $7.50 a year, plus $1 postage for all subscriptions outside the continental United States. Single copies, $1. Second class postage paid at Madison, Wisconsin, Copyright 1963 by Dembar Educational Research Services, Inc. Address all correspondence care of DERS, Box 1605, Madison 1, Wisconsin.

Published by **DEMBAR EDUCATIONAL RESEARCH SERVICES, Inc.** WALTER FRAUTSCHI, President. Prof. WILSON B. THIEDE, Vice President and Publisher. Prof. CLARENCE A. SCHOENFELD, Assistant to the Publisher. ARNOLD CAUCUTT, Treasurer and Business Manager. JEAN WILLIAMS POTTER, Editorial Services.

It Says in The JER ...

HOW VIABLE ARE the "natural sciences" "social sciences" and "humanities" as major classifications of scholarship? Not very, say Knapp and Ehlinger after building profiles based on the production of Ph.D.'s. Among the so-called "sciences," a major distinction has to be made between physical and biological sciences. Even among the biological sciences, those oriented to "plant" subject matter keep close company, but they are not significantly associated with "animal" disciplines. The homogeneity of the so-called "social sciences," in turn, would appear to be largely mythical. The "humanities," on the other hand, is fairly well vindicated as a category. The field of education seems to constitute a special department apart from the traditional fields of scholarship.

MIDDLE-CLASS CHILDREN anticipate assuming responsibilities earlier than do lower-class children, Zunich reports from a Texas study of 500 sixth-grade children.

TEACHERS CAN, following a period of in-service education, differentiate between pupils of varying mental health status, Bledsoe concludes after comparing teacher ratings with pupil performance on a mental health analysis.

HORN CALLS ATTENTION to multiple discriminant function analysis as a possible approach of one of the oldest problems in psychology — the discovery and measuring of traits, those relatively enduring characteristics which distinguish one person from another despite variation over a considerable range of circumstances.

ATTITUDE TOWARD MATHEMATICS is related to a broad constellation of personality variables indicative of adjustment and interest, Aiken reports from a study of college women. "Mathephiles" tend to be more socially and intellectually mature, more self-controlled and place more value on theoretical matters than do "mathephobes."

WHAT ABOUT SCHOOL FACULTY MEETINGS? Blumberg and Amidon say teachers are a lot happier with them when they are teacher-centered rather than principal-centered.

SANDERS WARNS that when teachers have access to intelligence test scores, their grading may be biased as they sub-consciously try to make their assessment of the student coincide with his primary mental ability.

DO PUPILS WHO SPEND an extra year in elementarv school come up to the level of regularly-promoted students? Kamii and Weikart looked at the marks, achievement levels, and intelligence test scores of Michigan seventh-graders retained once in elementary school, and found that they continue to rank below class average several years after.

THERE IS NO DIFFERENCE in the achievement of rural children, whether they are educated in a single-graded or multi-graded school unit. In other words, while a consolidated school may offer hot lunches and instrumental music, it does no better in terms of pupil achievement than the one-room school, at least according to a California study by Finley and Thompson.

A CHILD MAY ENCOUNTER DIFFICULTY throughout school as a result of entering the first grade before he is mature enough to do first-grade work, Dickinson and Larson suggest after studying 480 fourth-graders in South Dakota. Younger pupils with higher IQ's were not achieving as well as older children with lower IQ's, pointing to a "snowballing" effect under which differences that exist at early ages may become magnified as the group grows older.

SECTIONAL DIFFERENCES skew the performance of students on standardized tests, Campbell and Sheppard point out after administering the CAT to 2900 New York elementary pupils. They found national norms were not adequate as reference points for their school district. The development of local school district norms will aid in evaluating the performance of local students until a more valid test is found, they suggest.

TEACHING AIMED AT providing opportunity for all to learn may cause pupils of higher ability to halt progress while those of lesser ability are making larger gains. This is one of a number of interesting discoveries made by Scott in an extended study of 670 pupils in the Bend, Oregon schools.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 9, May - June 1963)

Marks, Achievement, and Intelligence of Seventh Graders Who Were Retained (Nonpromoted) Once in Elementary School

CONSTANCE K. KAMII
and DAVID P. WEIKART
University of Michigan

IN THE DESIRE for higher academic standards, attempts to raise standards are often made by homogeneous grouping and/or nonpromotion of those who do not attain certain levels of learning. This paper limits itself to the study of nonpromotion. It describes what kind of marks nonpromoted pupils are earning, how they score on achievement tests, and how they perform on a group test of intelligence several years after nonpromotion. This study is a descriptive one, but the ultimate question it attempts to answer is "Do children who spend seven years in elementary school do as well, in the long run, as those who finish it in six years?"

Procedure

The groups studied came from the seventh graders during the academic year 1959-60 in the two junior high schools in Ypsilanti, Michigan, a city of 21,500, approximately 20 miles west of Detroit. The general procedure was to describe a group of pupils who were retained once in elementary school and compare them with a sample of pupils who have never been retained.

The retained group was defined as the pupils in the seventh grade who had been in the Ypsilanti school system ever since kindergarten and who had a record of one (but not more than one) retention in grades one through five. Those who were retained in the sixth grade were not included, since this study did not concern itself with school performance immediately after retention. Those who had been retained more than once and those who were in special classes for the mentally retarded were also excluded, since these groups constitute extreme cases, and it was observed that they received mostly D's and F's in regular classes. Thirty-one pupils were found to fit the category of "retained." The various grades repeated were distributed as shown in Table I.

The normal group was defined as those seventh graders who had been in the Ypsilanti school system ever since kindergarten and who had never been retained and were in the expected age grade (i. e., those whose birthdays fall between December second, 1946 and December first, 1947). A random sample of 31 such youngsters was obtained.

The reason for drawing the two samples from those who came up through the Ypsilanti school system ever since kindergarten was to make the samples as comparable as possible and to exclude from the study youngsters who were retained in other school systems having a variety of promotion policies. It may be assumed that these pupils are from well-established families which do not move around, but it can also be argued, of course, that these families lack the initiative and aspiration which characterize mobility. It was assumed, at any rate, that factors such as mobility and socio-economic class were distributed randomly in both samples. Each group had 22 boys and nine girls.

Three variables were studied: 1) Marks received at the end of the first semester in the seventh grade in five academic subjects -- arithmetic, language, literature, science, and social studies; 2) achievement test scores in reading and arithmetic on the Iowa Every Pupil Tests of Basic Skills obtained toward the end of the sixth grade; and 3) intelligence test scores on the California Test of Mental Maturity. Marks in non-academic subjects were not considered because not everyone took art, music, etc.

Results

The marks in academic subjects which the 31 pupils in the retained group received two to six years after retention are shown in Table II, and Table III shows the percentage of each mark received.

The first thing that may be noticed is that approximately one-third of these marks are F's, one-third are C's, and the rest are mostly D's. When the D's and F's are combined, since both of these indicate unacceptable performance, the percentage

TABLE I

FREQUENCY DISTRIBUTION OF GRADES REPEATED

Grade Repeated	Number of Pupils
5	5
4	4
3	4
2	11
1	7
Total	31

TABLE III

MARKS EARNED BY 31 PUPILS IN THE RETAINED GROUP DURING FIRST SEMESTER IN SEVENTH GRADE

Mark	Frequency	Percentages
A	0	0
B	14	9
C	51	34
D	38	25
F	48	32
Total	151	100

TABLE IV

MARKS EARNED BY THE RETAINED AND NORMAL GROUPS DURING THE FIRST SEMESTER IN THE SEVENTH GRADE

	F	D	C	B	A	Total
Normal	0	16	52	67	17	152
Retained	48	38	51	14	0	151
Total	48	54	103	81	17	303

TABLE II

MARKS EARNED BY 31 PUPILS IN THE RETAINED GROUP DURING THE FIRST SEMESTER IN SEVENTH GRADE

Grade Repeated	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2	2	3	3	3	3	4	4	4	4	5	5	5	5	5
Sex	M	M	F	F	F	F	F	M	M	M	M	M	M	M	M	M	M	F	M	M	M	M	M	M	M	F	M	M	M	F	F
Subjects																															
Arithmetic	D	F	D	F	F	C	B	F	D	C	C	F	C	F	C	D	C	C	F	F	F	B	C	C	D	C	B	B	C	C	D
Language	D	F	F	D	D	C	C	F	F	C	D	D	D	F	D	F	C	C	F	F	F	C	D	C	D	C	B	C	C	B	D
Literature	D	F	F	F	F	C	C	F	F	C	-	-	C	F	C	D	C	C	F	D	F	C	F	C	D	C	B	C	D	-	F
Science	D	D	F	D	F	B	C	F	F	B	C	D	C	F	B	D	C	C	F	F	D	B	D	C	F	C	B	D	B	C	F
Social Studies	D	F	F	F	F	D	C	F	-	C	D	F	C	F	C	F	C	C	D	F	F	C	D	C	D	D	B	C	D	C	D

is approximately 57 percent. There is apparently little hope for anyone in this group to receive an A in an academic subject.

Let us for a moment turn our attention away from the group, take a closer look at the individuals, and see how many individuals in the retained group earned at least all C's except, to give them the benefit of the doubt, one D. Thirteen out of 31, or 42 percent, fit this category. The remaining 58 percent of the retained group earned D's and F's, mostly F's.

Table IV shows the comparison between the marks earned by the retained group and the normal group. The huge pile-up of F's in the retained group is a very sobering phenomenon, especially because the mark F is at the extreme of the distribution, and therefore its frequency is small in the over-all distribution of marks.

The statistical test used to determine whether or not the two groups could be considered to have come from the same population was the chi-square test. The null hypothesis tested was that the set of obtained frequencies differed no more than could be expected by chance alone from a set of expected frequencies on the basis that each mark could be expected to be distributed equally between the two samples. Chi-square came out to be 105.45, which is enormously greater than the value customarily required to reject the null hypothesis, and therefore it was concluded that the difference in marks which the two samples received was statistically significant beyond the one percent level of confidence. In other words, as far as marks in academic subjects are concerned, giving pupils an extra year in elementary school does not seem to make them come up to the level of the regularly promoted pupils whose chronological ages average a year younger. (Seventh-grade marks were selected as a sample of how once-retained pupils do several years later, and it was subsequently observed that their marks remained about the same during the second semester in the seventh grade and the first semester in the eighth grade.)

Achievement Test Scores

Since national norms are available on the Iowa Every Pupil Tests of Basic Skills, the two groups were compared with respect to these national norms. Tables V and VI show the distribution of reading-grade and arithmetic-grade levels achieved shortly before these seventh graders finished the sixth grade.

It is easily observed that the retained group as a whole achieved less in seven years than the normal group achieved in six years. In reading, only one-third (34 percent) of the retained group attained at least the sixth-grade level, while the corresponding percentage for the normal group was 84 percent. In arithmetic, 39 percent of the retained group attained at least the sixth-grade level, whereas 87 percent of the normal group achieved at least this level. Incidentally, the retained group in Tables V and VI has only 23 cases because eight in this group were absent on the day the test was given.

The t-test was used to determine whether or not these two samples could be considered to have come from the same population. The mean reading-grade level was 7.59 for the normal group and 5.67 for the retained group. The mean arithmetic-grade level was 7.37 for the normal group and 5.73 for the retained group. Since $\underline{t}$ came out to be 4.92 for reading and 4.91 for arithmetic, both of which are statistically significant beyond the one percent level of confidence, it was concluded with respect to achievement in reading and arithmetic that the retained sample who were "given" an extra year could not be considered to have come from the same population as the normal pupils.

Intelligence Test Scores

Table VII shows the distribution of IQ's by the California Test of Mental Maturity, and it was again observed that the normal group tested higher. The mean was 94 for the retained group and 112.6 for the normal group The t-test came out to be 5.38, which is statistically significant beyond the one percent level of confidence, and it was therefore concluded that, with respect to IQ's, the two groups could be regarded as having come from two distinct populations.

It was noted that less than half of the IQ's of the retained group (42 percent, calculated from the original, ungrouped data) were distributed in the range below 90, which is the lower limit of the range generally considered to be "average." The question that immediately came up was, "Did the pupils with below-average IQ's earn the D's and F's, which constituted 57 percent of all the marks earned by the retained group?" The next two sections represent an attempt to answer this question.

The Relationship Between Intelligence and Marks

Table VIII shows the relationship between IQ's and marks for the retained group. (The IQ intervals were set in such a way that the column totals would correspond proportionately to the row totals.) A positive relationship between the two variables was observed from this table, and the index of order association came out to be 0.58. (The corresponding index of order association for the normal group was 0.51.) Of significance is the fact that the D's and F's were not earned exclusively by the low-IQ pupils, and the C's were distributed throughout the entire IQ range. In short, low IQ's do not adequately account for the D's and F's earned by the retained group.

Table IX shows the relationship between IQ's and marks for the normal group. The IQ intervals in Table IX were set to correspond to those of

TABLE V

READING-GRADE LEVELS ACHIEVED BY THE RETAINED AND NORMAL GROUPS SHORTLY BEFORE THE END OF THE SIXTH GRADE

Grade Level	Retained	Normal
10.0 - 10.9		5 (16%)
9.0 - 9.9		3 (10%)
8.0 - 8.9		3 (10%)
7.0 - 7.9	1 (4%)	9 (29%)
6.0 - 6.9	7 (30%)	6 (19%)
5.0 - 5.9	10 (44%)	4 (13%)
4.0 - 4.9	4 (17%)	1 (3%)
3.0 - 3.9	1 (4%)	
Total	23 (99%)	31 (100%)

TABLE VI

ARITHMETIC-GRADE LEVELS ACHIEVED BY THE RETAINED AND NORMAL GROUPS SHORTLY BEFORE THE END OF THE SIXTH GRADE

Grade Level	Retained	Normal
9.0 - 9.9		2 (6%)
8.0 - 8.9		8 (26%)
7.0 - 7.9	4 (17%)	10 (32%)
6.0 - 6.9	5 (22%)	7 (23%)
5.0 - 5.9	8 (35%)	3 (10%)
4.0 - 4.9	3 (13%)	1 (3%)
3.0 - 3.9	3 (13%)	
Total	23 (100%)	31 (100%)

TABLE VII

IQ's OF THE RETAINED AND NORMAL GROUPS AS MEASURED BY THE CALIFORNIA TEST OF MENTAL MATURITY

IQ	Retained	Normal
125 - 134		4 (13%)
115 - 124	3 (11%)	12 (39%)
105 - 114	7 (25%)	6 (19%)
95 - 104	3 (11%)	8 (26%)
85 - 94	4 (14%)	1 (3%)
75 - 84	9 (32%)	
65 - 74	2 (7%)	
Total	28 (100%)	31 (100%)

TABLE VIII

RELATIONSHIP BETWEEN IQ's AND MARKS OF THE RETAINED GROUP

Mark	IQ Range					Total
	65-84	85-94	95-114	115-124	Above 125	
A						(0)
B		1	8	5		(14)
C	10	11	21	9		(51)
D	10	10	14	1		(35)
F	20	11	6			(37)
Total in IQ range	(8)	(7)	(10)	(3)	(0)	(137) (28)

TABLE IX

RELATIONSHIP BETWEEN IQ's AND MARKS OF THE NORMAL GROUP

Mark	IQ Range					Total
	65-84	85-94	95-114	115-124	Above 125	
A			6	6	5	(17)
B			21	37	9	(67)
C		3	32	13	4	(52)
D		2	10	3	1	(16)
F						(0)
Total in IQ range	(0)	(1)	(14)	(12)	(4)	(152) (31)

TABLE X

MARKS EARNED BY THE RETAINED AND NORMAL PUPILS WITH IQ's BETWEEN 90 AND 110

Group	Mark					Total
	A	B	C	D	F	
Retained	0	7	19	11	12	49
Normal	5	12	23	9	0	49

Table VIII (and no attempt was made to make the column totals correspond proportionately to the row totals.) The column totals in Table IX do not correspond to those of Table VIII, but it can be readily observed that, given the same IQ range, the regularly promoted pupils earned better marks than the ones who were once retained.

A more precise way of looking at the difference between the two groups with respect to the relationship between intelligence and marks is to hold IQ's constant and compare the marks earned by the two groups. Since the IQ range between 90 and 110 is generally considered to be average, the ten pupils in each group falling in this range were singled out, and their marks were compared (Table X). The chi-square test was used to test the null hypothesis that the set of obtained frequencies differed no more than a set of expected frequencies on the basis that each mark could be expected to be distributed equally between the two groups. Chi-square came out to be 18.88, which is greater than the value customarily required to reject the null hypothesis at the one percent level of confidence. The difference in marks earned by the two samples in the same IQ range was thus statistically significant, and the conclusion was again reached that, given the same IQ range, regularly promoted pupils earn better marks than the ones who have been retained. The lower IQ's of the retained pupils do not account very well for their poor marks.

Relationship Between Achievement Levels and Marks

If IQ's do not explain the poor marks earned by the retained group, the lack of certain basic skills might be a better explanation. For example, work in literature and social studies depends heavily on reading skills, and therefore a high correlation could be expected between reading achievement levels and marks in literature and social studies.

Table XI shows the relationship between reading achievement levels and marks in literature and social studies for the two groups. The retained group did not make a single A or B in these subjects, whereas the A's and B's earned by the normal group were spread over almost its entire range of reading skills. Also, given the same ability in reading, the regularly promoted pupils consistently earned better marks in these two subjects than did the retained ones.

Discussion

The purpose of this study was to find out what happens to retained youngsters several years after retention. The evidence indicates that, in general, retention does not help in the long run. This study thus went a step beyond Coffield's conclusion (1) that "There is no advantage in terms of seventh-grade achievement, for failure experienced at any particular grade level from grades three through six."

Table XII shows the marks received in arithmetic, reading, and spelling by 24 retained pupils immediately after retention. (Seven of the 31 pupils in the retained group are not included in this table because they received "satisfactory" and "unsatisfactory" instead of letter marks.) The marks in Table XII do not compare favorably with the marks received by the total elementary school population in Ypsilanti, which consist approximately of 26 percent A's, 46 percent B's, 20 percent C's, six percent D's, and two percent F's.

Learning requires both intelligence and motivational factors. As far as intelligence is concerned, the majority of low achievers have at least average intelligence, and some of them indeed have above-average intelligence. Goodlad (2) said, "The total body of evidence suggests the closer affiliation of undesirable social and personal adjustment characteristics with nonpromotion than with promotion. Although the exact causal nature of this affiliation cannot be ascertained with finality, there are clear indications that nonpromotion is the less defensible educational practice." Different teaching methods aimed at enhancing positive motivation and better social and personal adjustment would be a more effective solution than retention.

Ypsilantis and Bernert (3) reported that about four million pupils, or about one-fifth of the total school population, were retarded one year or more in their age-grade placement in the U. S. in 1950. They further discovered that youth who are retarded in their age-grade progress are most likely to drop out of school early. The problem of retention thus has far-reaching economic and social implications.

Whatever variable is studied, an overlap between the two groups is observed. This means that some retained pupils turn out to be better students than some regularly-promoted ones. Whether or not retention helped these students cannot be ascertained without a controlled experiment. Further research is necessary to determine what kinds of pupils benefit from retention before educators can justifiably retain pupils of elementary-school age.

Summary

The marks, achievement levels, and intelligence test scores of seventh graders who were retained (nonpromoted) once in elementary school were studied several years after retention, in an attempt to answer the question, "Do pupils who spent an extra year in elementary school come up to the level of regularly promoted ones?" The pupils who were retained once were found to have the following characteristics:

1. Their marks in academic subjects are signifi-

TABLE XI

RELATIONSHIP BETWEEN READING-GRADE LEVELS AND MARKS IN LITERATURE AND SOCIAL STUDIES

Group	Mark	Reading Grade Level								Total
		3.0-3.9	4.0-4.9	5.0-5.9	6.0-6.9	7.0-7.9	8.0-8.9	9.0-9.9	10.0-10.9	
Normal	A				1				2	(3)
	B			1	6	5	4	6	7	(29)
	C		2	3	5	12	1			(23)
	D			3		1	1			(5)
	F									(0)
	Total in range of reading level	(0)	(1)	(4)	(6)	(9)	(3)	(3)	(5)	(60) (31)
Retained	A									(0)
	B									(0)
	C		2	7	7	2				(18)
	D		2	9	1					(12)
	F	2	3	4	5					(14)
	Total in range of reading level	(1)	(4)	(10)	(7)	(1)	(0)	(0)	(0)	(44) (23)

TABLE XII

MARKS IN ARITHMETIC, READING, AND SPELLING EARNED BY 24 PUPILS THE SECOND TIME THROUGH THE SAME GRADE

Mark	Grade Repeated					All Grades
	1	2	3	4	5	
A	1	1	0	0	1	3 (5%)
B	3	6	2	1	7	19 (29%)
C	6	12	2	2	6	28 (42%)
D	2	6	2	3	1	14 (21%)
F	0	2	0	0	0	2 (3%)
Total	12	27	6	6	15	66 (100%)

cantly below the average of the regularly-promoted seventh graders, and the majority of their academic marks are D's and F's.

2. Their achievement levels in reading and arithmetic are significantly lower than those of the regularly-promoted pupils.

3. Their IQ's are significantly lower than those of the regularly-promoted pupils, but more than half of the retained pupils have at least average IQ 's as measured by the California Test of Mental Maturity.

4. The reason for getting D's and F's can be attributed neither to low intelligence nor to poor basic skills, such as reading. (However, indices of order association of the order of 0. 4 to 0. 6 were found between intelligence and marks and between achievement levels and marks.)

It was noted that the differences between the retained pupils and the regularly promoted ones would have been even more dramatic if those who had been retained more than once had been included in the retained group.

REFERENCES

1. Coffield, W. A. "A Longitudinal Study of the Effects of Non-Promotion on Educational Achievement in the Elementary School," Dissertation Abstracts, XIV, (1954), pp. 2291-92.

2. Goodlad, J. I. "Some Effects of Promotion and Non-Promotion Upon the Social and Personal Adjustment of Children," Journal of Experimental Education, XXII (1954), pp. 301-28.

3. Ypsilantis, J. N., and Bernert, Eleanor H. "Variation in Age-Grade School Performance," Teachers College Records, LVIII, (1957), pp. 268-77.

THE JOURNAL OF EXPERIMENTAL RESEARCH
(Volume 56, Number 9, May - June 1963)

The Discovery of Personality Traits

JOHN L. HORN
University of Denver

THIS NOTE ADDRESSES itself to one of the oldest problems in psychology, viz., the problem of discovering (or inventing, cf. Kelly, 1958) general measures of the relatively enduring characteristics which distinguish one person from another despite variation over a considerable range of circumstances. Trait theory supposes that there are such characteristics and numerous methodologies have been used to express such characteristics in terms of tests or inventories and the numbers which can be obtained by use of such devices. This paper proposes yet another methodological approach to the problem. It is not based on a new mathematical technique, but rather on a technique which has been neglected in psychology and education and one which, it seems to us, promises a more rigorous definition of the generic concept.

Two terms are intentionally left ambiguous in the verbal definition of traits given above. The terms "relatively enduring" and "over a considerable range of circumstances" are meant to imply that the temporal and situational lability of a supposed trait are not here to be stipulated, but must be related to one's particular theory of trait, which latter will stipulate the conditions under which the rank orders of persons on measures of the trait are supposed to remain stable. Present-day theories of trait differ in the modalities – e.g., ability, temperament, schemata, etc. – which they recognize, often implying that stability will vary with modality. They differ also in the emphasis they give various concepts which must be distinguished from trait, as, for example, role, state, the dynamic component of motive, etc. And they differ in many other ways.[2] Further research must furnish the basis for deciding between theories. Our purpose is merely to suggest some operations which will aid in this research, operations which seem particularly suited for the discovery of traits and for distinguishing the trait concept from other concepts used in personality theory.

I Existing Methods for Inferring Traits and Their Limitations

The method of factor analysis is often considered the *sine qua non* of trait discovery. As compared with item analysis, cluster analysis, and similar techniques used in the invention of a trait measure or the discovery of clusters among such measures, factor analysis is effectively used to help find essentially new ways of conceptualizing the dimensions of assessment – i.e., to discover traits. For this purpose, it is probably unrivaled among present-day techniques. But, as Humphreys (1952) has suggested, factor analysis is probably only one of the best methods for establishing some of the facts needed for proper inference of a trait, not the only method and probably not sufficient in itself. He lists four kinds of information which he believes should be garnered to support the claim for discovery (or invention) of a noteworthy trait, viz., 1) *homogeneity*, a necessary though not sufficient condition, evidence for which is given by an internal consistency coefficient for the behaviors of which the supposed trait is comprised, 2) *independence*, a test for which requires that the multiple correlation (perhaps regressed for sampling error) of a measure of the supposed trait with a large number of other, supposedly different traits, is less than unity when corrected for the lack of precision of the measures, 3) *stability*, and he adds, in a stable environment, evidence for which is supplied by test-retest ceofficients obtained over varying intervals of time and by knowledge of change in the environment, and 4) *validity*, by which is meant in this context (a) significant correlation[3] of the measure with practical criteria (i.e. there are potentially several validities) and/or (b) demonstration that explanations for the observed behavior of the traits are consistent with subsequently obtained correlations which the measure of the trait has with other measures.

A factor analysis is probably most often used, and is best suited for, structuring the data to obtain evidence for criteria (1) and (4). Requiring that vectors not approach each other too closely in the rotation process (and this in turn will usually require that not too many factors be extracted) tends to insure (2) within a given analysis. Also, (2) is possibly a special case of (4), in which case Cattell's (1957) emphasis (when interpreting factors) on examining what is not loaded by the factor, as well as what is loaded, is an approach via the factor analysis to meeting this criterion. Criterion (4b), relating to what has become known as “construct validation”, needs to be applied gingerly, of course, because consideration of sample size and other factors almost always leaves room for considerable debate concerning what is and what is not a “proper” amount of “correlation”. Our main concern here, however, is with criterion (3).

An R-technique factor analysis establishes that behaviors occur together within a person (on an average computed across persons) on a given occasion. In itself it provides no evidence whatsoever about the stability of the relationship or about the stability of the rank ordering of people for the amount of influence of whatever is inferred from the relationship. A repeat analysis with some of the same variables will provide evidence for the stability of the relationship, but does not bear on the question of the stability of rank ordering (even if the same sample of people is used in both analyses, provided, of course, stability coefficients are not involved). It must be granted that some kinds of data — e.g. ratings for the way persons characteristically behave or ability test scores — permit development of a fairly strong case for the contention that rank ordering will remain stable over fairly extended periods of time, with noteworthy changes of global situation, etc. But even many of these cases must depend rather a great deal on “common sense”, on equally fallible incomplete observations, or on certain other kinds of external evidence. Increased application of factor analysis to the problems of motivation and group behavior has made it apparent (if it were not before) that a factor among measures of individuals differences which is highly replicatable, and is therefore sometimes treated as implying stable trait, may not permit a rank ordering of individuals that will remain stable over occasions. For example, one can probably take a number of measures of thirst, compare individual differences on these in almost any sample of persons he might care to select, and get a very similar factor in each sample. But we wouldn't wish to argue — or imply — that the rank order of people on a sum of these measures will remain stable for very long. Cattell (1957) uses the term “state” to refer to such functional unities. There are other examples where common sense is less helpful in deciding whether one has isolated a trait or a state or something which deserves another name. Would mutual ratings for “sociability” by fraternity brothers be expected to lead to a rank ordering of these individuals that would correlate substantially with a rank ordering based on equally reliable ratings[4] (obtained from different raters) made when these men were in basic training during their hitch in the service? “Maybe.” Perhaps even “probably.” But the question posed by examples like this have led some theorists (e.g. Newcomb, 1949) to propose that role, not trait, is the most reasonable kind of concept to use for the conceptualization of the major portion of what is studied under the heading of personality. Similarly, would the rank ordering obtained on Saturday night for these same fraternity men, using measures from a replicatable factor interpreted as the mating erg (cf. Cattell, 1957) be expected to correlate highly with the ordering obtained on Sunday morning? Again the answer must be “maybe” or “probably”, your answer, depending on your conception of what is measured in this factor and your conception of the way fraternity men behave on Saturday nights. But the point is that the term trait is often applied in studies where not enough information has been gathered to determine the answers to the questions implied by this usage, and where the analysis has not been designed to permit answer to these questions.

In sum, then, R-technique factors — however they are interpreted — may suggest measures for reliably distinguishing one person from another over fairly extended periods of time, over changes of global situations, etc., or they may not, and instead suggest measures for distinguishing the person on one occasion from the same person on another occasion. More likely, perhaps, such factors may suggest some of the measures which can be used to characterize both kinds of individual differences. Cattell and Scheier (1958) have suggested, for example, that some of the measures which define the R-technique factor of anxiety are highly susceptible to change by a particular set of conditions designed to induce anxiety, while other measures in this factor are a good deal less susceptible. The rank order of individuals may thus alter with these conditions. Perhaps it is not unreasonable to suppose that measures which are susceptible to change under one set of conditions will remain stable, while other components change, when another set of conditions are imposed. The rank order of individuals may again change and may be different than that in the former situation.

Several methods have been proposed for the purpose of distinguishing traits, states, etc. Incremental R-technique factorization (Cattell and Scheier, 1958) is one such method. In our opinion this is not a good method for distinguishing the stable from the labile components of a set of measures, largely because both kinds of factors can appear in the solution, either kind can tend to cancel the other kind out, and it is not unequivocally shown whether a factor is mainly due to covariation on a given occasion, to covariation over occasions, or both. This is illustrated by consideration of the coefficient upon which the incremental factorization is based. This coefficient is the correlation between two difference scores, viz., the score on test X on occasion one minus that on occasion two, and the score on test Y on occasion one minus that on occasion two. The coefficient is then written:

$$(1)\quad r_{(x_1 - x_2)(y_1 - y_2)} = \frac{r_{x_1y_1}S_{x_1}S_{y_1} + r_{x_2y_2}S_{x_2}S_{x_2} - r_{x_1y_2}S_{x_1}S_{y_2} - r_{x_2y_1}S_{x_2}S_{y_1}}{\sqrt{(S_{x_1} + S_{x_2} - 2r_{x_1x_2}S_{x_1}S_{x_2})(S_{y_1} + S_{y_2} - 2r_{y_1y_2}S_{y_1}S_{y_2})}}$$

It will be seen that if the correlation between two tests on each occasion is high, but the correlation between tests over occasions is low – i.e., the stability within the person is low – the coefficient will mainly reflect "state". If the converse is true – i.e., stability is relatively high – the coefficient will either tend to zero – when the correlation on a given occasion is about the same as that across occasions – or will tend to be negative. In the latter case the measures would show more stability than internal consistency. This would furnish some basis for inferring trait. However, one couldn't tell by looking at the factor whether to infer trait or state. It could represent either kind of covariation. Also, the possibility of the coefficient tending to zero should not be underestimated, for in many applications it seems likely that the correlation on a given occasion would be of about the same order, and have the same sign, as that across occasions.

The factorization of correlations between estimated true gain scores (Lord, 1956; Lord, 1958; McNemar, 1958) seems best suited for the discovery of factors-in-change. Here the trait component is suppressed. The person who scores high, say, on a set of tests on two occasions, but does not change much, will tend to have a lower gain score than the person who scores low on one occasion and high on the other. The unity represented by a factor in a set of such scores represents the performances that are changing together. In theory, the rank order of scores for a set of persons might remain the same. The factor would then indicate that everyone had changed, some more than others. This would seem a very useful method for isolating performance improvement resulting from educational exposure, but it does not center on the problem of isolating stable individual differences, which is the problem we are concerned with in this paper.

A search for simplex matrices in retest data (cf. Humphreys, 1960) aims at the discovery that a particular kind of systematic change in rank order of people on a given set of measures is occurring. A simplex in retests with measures of intelligence suggests, for example, that some people are moving up in their relative standing on amount of intelligence, while others are moving down. This kind of demonstration can also be very useful in suggesting or verifying theories – e.g., those specifying differences in rates of maturation. But again, it does not bear directly on the problem posed in this paper.

The factorization of simplex matrices (DuBois, 1960) would seem to be mainly an aid for inferring the order of change for variables. But we have not really looked into this matter, and therefore do not propose to push this argument here.

II A Method Which Aims to Overcome Some of the Limitations of Existing Methods.

Consider that one has a set of measures representing retests of the same N people (1, 2, . . . p, . . . N) on the same or parallel forms of m tests (1, 2, . . . , i, j, . . . , m) given on k occasions (1, 2, . . . t, . . . k). Consider the sum of cross-products for tests i and j over k occasions for a given person, i.e., the sum that would be obtained in a P-technique factorization. A "within" cross-products term for tests i and j over all N persons would then take the form (2)

$$(2)\quad w_{ij} = \sum_{p=1}^{N} \sum_{t=1}^{k} (X_{ipt} - \overline{X}_{ip})(X_{jpt} - \overline{X}_{jp})$$

(and an estimate of the within covariance could be obtained by dividing w_{ij} by the appropriate degrees of freedom). Designate the symmetric matrix containing w_{ij}'s for all m tests as "W". This $m \times m$ matrix would have rank $\leqslant$ m.

For convenience we set the total mean for tests j (j=1, . . . , m), over occasions and people, at zero. In general the mean for person p on test j over k occasions is not thereby set equal to zero, however. In fact, this score is an indicant of this person's "typical" level of performance. A cross-product term comparable to w_{ij} above, but based on this "typical" score would then take the form

$$(3)\quad b_{ij} = m \sum_{p=1}^{N} \overline{X}_{ip}\,\overline{X}_{jp}$$

where the assumption that the total mean (over t and p) for each variable is zero makes it unnecessary to write the person means as deviations from these totals. Designate the symmetric matrix containing b_{ij}'s for all m tests as B. This $m \times m$ matrix would again have maximum rank $\leqslant$ m.

Now if one wishes to infer trait, he is by our arguments mainly interested in inter-individual differences in "typical" level, but he must view this relative to the intra-person variation which results from change of occasion. One possible way to state this mathematically is to ask "What is the maximum ratio of between-persons to within-the-person variation?" With this as the question the task is to find a vector $V' = [v_1, \ldots, v_m]$ (or, actually, the several vectors) such that

$$(4) \quad \frac{V'BV}{V'WV} = \lambda$$

is maximized. Taking the partial derivatives with respect to $v_1, \ldots, v_m$, and setting the result equal to zero

$$(5) \quad \frac{(V'WV)BV - (V_1BV)WV}{(V'WV)^2} = 0$$

using the identity in (3) and multiplying through by $(V'WV)$, and we have

$$(6) \quad BV - \lambda WV = 0$$

If we assume that W^{-1} exists[5]

$$(7) \quad (W^{-1}B - \lambda I)V = 0$$

which has nontrivial solutions if and only if

$$(8) \quad (W^{-1}B - \lambda I) = 0$$

where, for each nontrivial and distinct characteristic root, $1, 2, \ldots, f$ ($f \leqslant m$), there is an associated nonzero vector. Analogous to the largest latent root in an ordinary R-technique factorization, the first latent root here suggests the variance of the factor which can, in this sample of tests and people, most effectively distinguish between people despite variations of the situation. Successive latent roots "account for" progressively less of this between persons variance. To test the hypothesis specifying no difference between people, the statistic

$$(10) \quad \frac{(Nk - m - 1)}{m(Nk - 2)} \lambda$$

can be used as the variance ratio with m and (Nk - m - 1) degrees of freedom (Rao, 1952). Thus, the number of factors can be estimated at some specified level of confidence. These suggest the stable traits in a given set of people obtained on a given set of measures and over a given set of occasions.

The set of vectors, $V_1, V_2, \ldots, V_f$, corresponding to the significant latent roots, constitute a basis, F_1, for this system. This needn't be the psychologically most meaningful basis. A more meaningful basis, maintaining the dimension of the vector space, might be obtained by choosing a nonsingular matrix, T, that would transform F_1 into another set of vectors, F_2, having certain desired properties – e.g., those of simple structure. The logic of simple structure was originally developed (Thurstone, 1940) within the context of trait theory, and seems appropriate in the present context for many of the same reasons that Thurstone advanced. The question is partly an empirical one, however, and cannot be answered with finality at the present time.

III Anticipation of Some Objections to the Proposed Method

As was noted in our introductory remarks, the mathematical techniques here employed are not by any means new. Multiple discriminant function analysis has been used in the biological and social sciences for a number of years (e.g., see Rao, 1952). The particular application of discriminat function analysis which is here proposed is novel in the sense that it has not, to our knowledge, been tried, or its implications fully explored, but the idea of such application was put forward tangentially by Tiedman (1951) some ten years ago. Yet few investigators would seem to be aware of the mathematical technique, and fewer still would seem to have considered the application we have proposed. Thus, our purpose has been mainly to call these matters to their attention. The method no doubt has several mathematical, logical and practical shortcomings. We shall not dwell on these, for their true impact cannot be accurately guaged until the method is tried, but we shall mention a few of the more obvious difficulties likely to be encountered in applications of the method.

A least squares computation always implies some shrinkage with change of sample. In this respect the present method is no different than many other mathematical techniques in use in the behavioral sciences. Indeed, mathematically the method is based on assumptions that are very similar to those involved in traditional factor analysis. Assumptions will no doubt be violated in actual research. Whether or not these violations will work a serious hardship on interpretation of results, is a question that cannot be answered without further study.

Because the method requires the collection of retest data, one might object to it on grounds that it is not, shall we say, easy to apply. This, of course, is hardly sufficient grounds for rejection of the method in principle, however much it may justify rejection on grounds of practical feasibility. In adopting this policy one must be aware that most theories in the personality field make quite a verbal distinction between stable individual differences and individual differences which reflect different amounts of exposure to changing global circumstances. We cannot hope to provide an operational basis for such theoretical niceties, and a rational basis for choice between theories, without employing methods which will allow empirical demonstration of the distinctions drawn.

A slightly related practical objection is that the method appears to require either retest with the same tests, leading to several difficulties concerning how to interpret any stability or lack of stability which may be observed, or to require parallel forms of tests, which latter are of well-

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 9, May - June 1963)

Teacher Reactions to School Faculty Meetings

ARTHUR BLUMBERG and EDMUND AMIDON
Temple University

THE LITERATURE of school supervision and administration leaves little doubt about the importance of faculty meetings as they relate to the total educational program. For example, Reeder states that "The teachers' meeting . . . assumes the most important role in our plan for supervision."[2] The reasoning suggested by this point of view is that supervision is, for the most part, a process of group leadership and that the one chance where there is an opportunity for this leadership to be exercised most broadly and effectively is when the entire faculty is gathered together.

Specific functions of school faculty meetings are summarized by Wiles in the following manner:

> They are described as opportunities for cooperative thinking, for staff planning, for the presentation of stimulation talks by resource people, for getting to know the total school, and for interchange of ideas -- all of which result in growth for the staff member.[3]

Implicit in this concept is the thought that, in addition to the potential for enhancing an individual teacher's personal and professional growth, faculty meetings can also aid in the establishment and maintenance of sound school organization and high staff morale.

It does not take a keen observer, however, to discover that, at least on the face of it, these potentials are probably not being met or, in some cases, even approached. Rarely does one find a teacher who is enthusiastic about the prospect of attending these meetings. They are more often seen as "all right," "a necessary evil," or "a downright waste of time."

Why should there be such a wide divergence between what is theoretically possible to accomplish in these situations and what it appears from casual observation most teachers perceive as being accomplished? Reasons frequently given for this are that the meetings are not held on school time, that they last too long, or that the facilities are not conducive to productivity; but these points are mechanical in nature. While it is possible that, at times, they are crucial in their effect on outcomes, it seems much more likely that the relative success or failure of faculty meetings is primarily dependent on the dynamics of the situation in question. That is, if one is interested in inquiring into faculty-meeting problems in a productive fashion, the approach should be behaviorally oriented rather than mechanically oriented.

Published research that is concerned with school faculty meetings is extremely limited. Though it is possible that other studies are available, our review of educational research journals turned up only two investigations that seemed to bear on the problem. Interestingly enough, these investigations were concerned primarily with the effectiveness of curriculum study meetings and not with what might be conceived of as an "all faculty" group that meets to consider general school problems.

One investigation conducted by the Ohio Education Association[4] was aimed at factors that are perceived by teachers as being related to their satisfaction with faculty meetings. As might be expected, the respondents enumerated such things as irrelevant discussions, unplanned agenda, and leader domination as affecting their reactions negatively.

A study conducted by Smith[5] focused directly on the status-leader role in these meetings. Effective leader behavior was seen as that directed towards clarifying, reflecting member feeling, calling attention to available resources, and sensitizing members to group work methods. When these behaviors were engaged in by the leader, it was found that there was increase in member responsibility, participation, volunteering for jobs, and completing tasks. These findings seem to be completely in line with numerous other researches that have been concerned with the kinds of leader behaviors that are positively related to effected group problem-solving.

The study described in this article focuses on some of the behavioral and attitudinal factors that may be related to the way teachers react to their faculty meetings and to the manner in which they perceive the quality of the interpersonal relationships that exist in their school. The study was concerned with two major independent variables: 1) the kind of faculty meeting (based on different activities), and 2) the types of schools, elementary or secondary. The purpose of this design was to determine the relationships that might exist between these variables and the satisfaction teachers felt with their meetings as well as attitudes they might hold concerning the school and the people who work in it. In a sense, perhaps, the central question of the study is, if we characterize the faculty meeting, can we also describe teacher satisfaction with these meetings and other, more general, attitudes they may hold about the school and the faculty?

Procedures

Data were collected for the study by administering a questionnaire to 40 elementary and 49 secondary in-service teachers enrolled in an introductory group dynamics course during the Spring semester, 1960, at the College of Education, Temple University.

In order to determine the kind of faculty meeting that each teacher had in his school, the initial question in the instrument was designed to find out to what extent certain activities took place in each teacher's faculty meeting. Activities used to describe faculty meetings in this item included principal giving opinions and information, principal reprimanding teachers, teachers giving opinions and information, principal and teachers discussing problems, and teachers reporting on problems. Teachers were asked to describe their meetings by marking "never," "once or twice," "rather often," or "very often" for each of the activities listed.

In this study, the only questionnaires analyzed were those on which the teachers indicated, by the kinds of activities that took place in their meetings, that their faculty meetings were either very principal-centered or very faculty-centered. Principal-centered meetings were characterized by much principal-initiated activity and little activity initiated by the faculty. The faculty-centered meetings were those in which the teachers initiated most of the activity and the principal relatively little.

Using the above criteria, we selected for analysis the 25 percent of the teachers who described their meetings as being most faculty-centered and the 25 percent who saw their meetings as being most principal-centered. Each of these groups was composed of half elementary and half secondary teachers.

These two classifications made it possible to make comparisons between faculty-centered and principal-centered groups and between groups of elementary and secondary school teachers. The following four subgroups were also identified for analysis:

1. Faculty-centered, elementary school . . . 10 teachers
2. Faculty-centered, secondary school . . . 10 teachers
3. Principal-centered, elementary school . . . 10 teachers
4. Principal-centered, secondary school . . . 10 teachers

Results of the Study

A casual look at the answers on the satisfaction question indicated that, in general, no groups of teachers in the study felt that faculty meetings were an effective use of time and energy. Answers to this question ranged from "faculty meetings were fairly satisfactory" to "faculty meetings were a waste of time." Only six teachers out of the 40 included in the analysis thought that their faculty meetings were an effective use of time and energy, while 28 thought faculty meetings were a waste of time.

The first step in the formal analysis of data was to run tests of significance between the two types of schools and the types of faculty meetings on the major variable of satisfaction as well as attitudes and perceptions teachers held toward faculty and the principal. The analysis of variance technique was used to make these comparisons.

In the initial analysis, differences significant at the .01 level of probability were found between the faculty-centered groups and principal-centered groups on level of satisfaction, the extent to which teachers felt their faculty was alert or bogged down, the degree of closeness to fellow faculty members, the amount of freedom teachers felt other teachers had to speak in meetings, the amount of freedom a teacher himself felt he had to speak in meetings, and the extent to which a teacher's preferred pattern of activities actually took place in faculty meetings.

There were no significant differences at or above the .05 level of probability between the principal-centered and the faculty-centered groups on the variables of difficulty in talking to the principal, perceptions of the principal as being critical, or the extent to which the teachers felt free to speak to the principal outside the faculty meeting.

In one sense, the significant differences were all in favor of the faculty-centered group. That is, these teachers were more satisfied, felt their faculty more alert, thought the faculty members were closer, saw their principal as being less critical, felt the teachers had more freedom to speak, and perceived less conflict between their preferences and what actually happened in the meetings than did the principal-centered group.

There was a significant difference between the elementary and the secondary teachers on the item concerned with the amount of freedom teachers felt to speak in faculty meetings. The elementary teachers felt freer to speak in their faculty meetings than did the secondary teachers. There was also a significant difference between the elementary and secondary groups on size of faculty. The average size of the elementary faculty groups was about 20, while the average size of the secondary faculty groups was about 52. All these differences were significant at the .01 level of probability.

The second major step in analysis consisted in running t-tests between subgroups where the interaction F ratio on the analysis of variance was significant at at least the .05 level (faculty-centered elementary, faculty-centered secondary, principal-centered elementary, and principal-centered secondary). The results of these analyses showed that there was only one .05 or .01 level difference between the means of the faculty-centered secondary group and the principal-centered secondary group, in which the means of the two elementary groups were not significantly different from one another. This significant t-ratio was found on the item which was concerned with freedom teachers thought other teachers felt to express themselves. However, an inspection of the means in Table I indicates that the mean differences between the two high-school subgroups are consistently greater than the differences between the two elementary subgroups. Although only this one significant difference was found, there seems to be a discernible trend which indicates that there is a greater difference between the secondary teacher groups than between the elementary teacher groups on almost all of the items in the questionnaire.

The third step in the analysis was to rerun certain tests of significance, where high correlations between two variables warranted the statistical control of a particular variable on another variable. Therefore, several analyses of covariance were run where it was thought that further analysis would clarify some of the relationships which had already been identified.

The variables controlled were 1) the extent to which the teachers' preferred pattern of faculty meetings is actually accomplished, 2) the extent to which there was a conflict between the teachers' preferred pattern of faculty meetings and what actually occurred, and 3) size of faculty. In only two cases did the analysis of covariance seem to effect differences between groups where the analysis of variance had yielded significant differences. One was in the case of satisfaction when conflict between a teacher's preferred pattern and what actually happened was controlled, and the other was when size was controlled on freedom teachers felt to express themselves. The analysis of variance for satisfaction yielded an F-ratio significant at the .01 level when a test was made between the principal and faculty-centered groups; however, when the preference conflict variable was controlled, this F-ratio was reduced to nearly zero. When size was controlled on freedom teachers felt to express themselves, there was no longer a significant difference between elementary and secondary teachers on this variable.

Summary and Conclusions

The results of this study seem to confirm some of the assumptions made concerning 1) the relationships that may exist between the nature of specific group interaction and the kinds of interpersonal perceptions and attitudes that may exist concurrently within an organization, 2) the effects on satisfaction with an experience when one's preferences in regard to what ought to happen are continually denied, and 3) the effect of size of group on apparent freedom to interact verbally.

Our findings, for example, indicated that when teachers saw their faculty meetings as being relatively more faculty-centered (the interaction was more widespread with greater teacher responsibility being assumed) than principal-centered (the principal exercises almost complete direction and control), there were significant differences in regard to:

1. the degree of satisfaction with the meeting
2. teacher perceptions of the general feeling tone of the faculty
3. teacher perceptions of the state of interpersonal relationships that exist among the faculty
4. teacher perceptions of the extent to which the principal reacts critically to teacher behavior in faculty meetings
5. teacher perceptions of the freedom of other teachers to speak in faculty meetings
6. the degree of conflict between one's preference about what should happen in meetings and what actually does happen

In the case of all these variables, the faculty-centered meetings seemed to produce significantly more favorable reactions from teachers than did those situations which were principal-centered.

When comparisons were made between the two school groups, significant differences were found on only one variable. Regardless of the nature of their faculty meetings, elementary teachers saw themselves as feeling more free to speak in the meetings than did secondary teachers. (It is clear, however, that this difference can be accounted for by the relatively smaller size of elementary school faculties.)

TABLE I

MEANS OF GROUPS AND SUBGROUPS

Means of Groups	Satisfaction with faculty meetings	Attitudes toward the faculty in general	Relationship among members of the faculty	Relationship between the teachers and principals	Teachers' perceptions of principals' reaction to teacher ideas	Teachers' perceptions of other teachers' reactions to teacher ideas	Freedom teachers feel to express themselves	Freedom teachers feel other teachers have to express themselves	Agreement between what teachers want and what happens in faculty meetings	Conflict between what teachers want and what happens in faculty meetings	Size of faculties
Principal-centered Groups	2.10**	3.55**	4.30**	4.90	4.25	4.20	4.70	4.20**	3.40**	6.60**	3.75
Faculty-centered Groups	4.50**	6.05**	6.20**	6.15	6.55**	5.00	5.00	6.30**	4.90**	4.00**	3.30
High School Group	3.35	5.00	5.50	5.30	5.45	4.20	3.90**	5.20	4.20	5.80	4.80**
Elementary School Group	3.25	4.60	5.00	5.75	5.35	5.00	5.85**	5.30	4.00	5.00	2.30**
Means of Subgroups											
Principal-centered Elementary Groups	2.50	3.60	4.10	5.20	4.40	4.50	6.00	5.00	3.30	5.80	2.90
Principal-centered Secondary Groups	1.70	3.50	4.50	4.60	4.10	3.90	3.40	3.44**	3.50	7.50	4.60
Faculty-centered Elementary Groups	4.00	5.70	5.90	6.40	6.40	5.40	5.70	5.70	4.70	4.10	1.70
Faculty-centered Secondary Groups	5.00	6.40	6.50	5.90	6.70	4.70	4.40	6.80**	5.10	4.00	5.00

** Means are significantly different at .01 level

In regard to perceived freedom of other teachers to speak in their faculty meetings, while no differences were found between type of meeting (faculty- or principal-centered) in the elementary situation, a significant relationship was indicated on this variable by secondary teachers. Those who perceived their meetings as faculty-centered also felt teachers were more free to speak out. It is interesting to note that when school-school comparisons were made on all the other variables, although no statistically significant differences were found, there was a consistent trend in this direction. That is, there was less difference between the reactions of elementary teachers regardless of the kind of meeting, but a definite, though not significant, difference in secondary schools dependent on the nature of the meeting.

A final point of summary is worthy of note. If the potential effect of conflict between individual preferences for faculty meeting activities and what actually occurred was eliminated statistically, there was no longer any difference between faculty- and principal-centered groups in the satisfaction level of the teachers. In other words, the satisfaction level is not affected by the activities of the meeting (faculty- or principal-centered) when the preference conflict is the same for both types of meetings. Interestingly enough, the results indicated that faculty-centered meetings could be characterized by less preference-conflict than those which were principal-centered. This seems to indicate that minimum preference-conflict may be a characteristic of faculty-centered meetings; and that it may be this characteristic which makes the faculty-centered meetings so much more satisfying to teachers.

Although this study shed some initial light about the nature of school faculty meeting dynamics, we would hesitate to make any generalizable conclusions about the way in which a particular kind of faculty meeting affected faculty attitudes and relationships in a specific school. Rather, what seems most important is to use the insights obtained from it as a basis upon which to conduct further inquiry into questions concerning faculty meeting productivity as it may be related to teacher morale, working relationships, and individual and professional growth.

FOOTNOTES

1. This study is part of an ongoing research program in educational problems that is being conducted by the Group Dynamics Center, Temple University.
2. Edwin H. Reeder, Supervision in the Elementary School (Boston: Houghton Mifflin Co., 1953), p. 211.
3. Kimball Wiles, Supervision for Better Schools (New York: Prentice-Hall, Inc., 1950), p. 150.
4. D. L. Arnold, "What Makes Teachers' Meetings Satisfying and Effective?" Ohio Schools, LX (March 1952), p. 293.
5. Mary Smith, "Action Research to Improve Teacher Planning Meetings," School Review, LX (March 1952), p. 146.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 9, May - June 1963)

A Comparison of the Achievement of Multi-Graded and Single-Graded Rural Elementary School Children

CARMEN J. FINLEY and JACK M. THOMPSON
Sonoma County Schools, California

SCHOOL DISTRICT reorganization in California and elsewhere in the nation progresses slowly. It has been shown that there is an optimal size range which promotes efficient administration and a good instructional program. Support for this program of reorganization can be demonstrated on the basis of a lower unit cost and added services such as transportation, hot lunches, instrumental music, nursing services and the like. Are these advantages reflected also in higher achievement for children in larger schools? This is the question with which this paper is concerned.

It is recognized that this particular issue of achievement can only be narrowly investigated with the present available testing instruments. However, with current interest in the ungraded primary class or multi-graded grouping in larger schools, the question arises as to whether the one- and two-room schoolhouse may not have been forced into a desirable curriculum grouping years ahead of its time.

The hypothesis of this study is that there is no difference in the achievement in basic subjects of rural children in multi-graded classrooms as compared with children attending single-graded classrooms.

A review of the research studies in this area shows conflicting results. Drier (3), Adams (1), McIntosh and Schrammel (7) and Nelson (8) found no significant differences in achievement between children educated in single-graded and multi-graded classes. Martens (6), Foote (4), Clem and Hovey (2) and Wilson and Ashbaugh (9) found differences favoring children educated in a graded school. On the other hand, Hull (5), in evaluating multi-graded teaching within a graded school system, found differences favoring the multi-graded group.

Some of the above studies have certain limitations. McIntosh and Schrammel (7) and Foote (4) used no tests of significance. McIntosh and Schrammel (7), Clem and Hovey (2) and Nelson (8) did not equate the multi-graded and single-graded groups. Hull's (5) conclusion of higher achievement for the multi-graded group might be questioned as statistically significant differences were found for only five of the 18 tests made in achievement areas.

Method

Grades three and five were chosen for investigation. Both multi-graded and fully-graded populations consisted of all pupils at these grade levels in Sonoma County, California. The third graders must have been in attendance at school the two previous years (as first and second graders). The fifth graders must have been in attendance at the school the three previous years (as second, third, and fourth graders). For the purpose of this study, the multi-graded school was defined as a school with four teachers or less.

The fully-graded population consisted of children attending a school large enough to have at least one teacher per grade level. The size of the fully-graded schools ranged from 300 to 1,322 average daily attendance.

The two groups were matched on the following basis:

1. Sex
2. I Q within five points
3. Chronological age within three months
4. Participation in the yearly county-wide group testing program. Since the fall testing program usually extends from September through December, it was further required that a given pair must have taken the achievement test within a month's time of each other (differences in opening dates of schools accounted for). Each member of the pair also must have taken both the

achievement test and the mental ability test during the current year.

The mental ability test used was the California Short Form Test of Mental Maturity. The achievement test used was the California Achievement Battery Form W.

After children were matched on the above criteria, their achievement records were investigated for the following test areas:

1. reading vocabulary
2. reading comprehension
3. arithmetic fundamentals
4. arithmetic reasoning
5. mechanics of English
6. spelling
7. battery

Mean differences and t-values were computed for boys, girls and total group. The five percent level was defined as indicative of statistical significance.

Subjects

The subjects for this investigation consisted of 53 paired boys and 51 paired girls at third-grade level and 62 paired boys and 46 paired girls at fifth-grade level who met the criteria for matching. The multi-graded group was drawn from 28 schools with no more than 12 of the total 212 coming from any one school. The fully-graded group was drawn from 23 different schools with no more than 21 coming from any one school. Table I shows the ability level and the chronological ages of the groups.

Results

Table II summarizes the results.

Third Grade: Boys. It may be seen that boys in fully-graded classrooms in larger schools tend to score better, on the average, in reading vocabulary, mechanics of English and spelling, while boys in smaller multi-graded schools score better, on the average, in reading comprehension, arithmetic reasoning, arithmetic fundamentals and on the composite battery score. However, none of the differences are large enough to be considered statistically significant. The largest t-value was found in arithmetic fundamentals with a probability of a difference this great occurring by chance between five and ten times out of a hundred.

Third Grade: Girls. As may be readily seen, there is a consistent tendency for girls in the larger, fully-graded classrooms to score higher, on the average, than girls in small, multi-graded situations. With the one exception of arithmetic fundamentals, all differences favor girls in fully-graded schools. However, none of the differences were large enough to be considered statistically significant. The one area in which multi-graded girls excelled fully-graded girls, arithmetic fundamentals, produced the highest t-value and most nearly approached statistical significance. The probability of obtaining a difference this large by chance is between five percent and ten percent. The only other area which approached a difference this great was spelling.

Third Grade: Total Group. When boys and girls are considered together, the differences found in reading comprehension, arithmetic reasoning and total battery tend to balance out. However, in the case of arithmetic fundamentals where the direction of the difference is the same for both sexes, one supports the other to produce a difference which is significantly in favor of third graders in the multi-graded school. The remaining areas of reading vocabulary, mechanics of grammar and spelling tend in the direction of favoring third graders in fully-graded situations. However, the differences are not great enough to be considered statistically significant.

Fifth Grade: Boys. As in the case of third-grade boys, fifth-grade boys tend to score better in reading vocabulary, although the difference is not great enough to be significant in either case.

For reading comprehension, the situation is reversed. While third-grade multi-graded boys tend to score better than their matched pairs, fully-graded fifth-grade boys tend to score higher. However, again, neither difference is great enough to be significant.

The most interesting situation is seen in arithmetic fundamentals. For fifth-grade boys, this is the only area in which a significant difference is found, the direction favoring boys in fully-graded schools. At the third-grade level, the difference was in favor of multi-graded boys, and although the difference was not significant, it was the one of the two subtest areas which most nearly approached significance.

Other differences for fifth-grade boys are negligible.

Fifth Grade: Girls. Fifth-grade girls show negligible differences for the most part, the highest t-value being found in reading vocabulary in favor of fully-graded girls. However, the difference is non-significant.

Fifth Grade: Total Group. When fifth-grade boys and girls are treated as a group, no significant differences are found. However, the subtest which most nearly approaches significance is arithmetic fundamentals and is in favor of fully-graded children.

TABLE I

ABILITY AND CHRONOLOGICAL AGES OF FULLY-GRADED AND MULTI-GRADED GROUPS

	I. Q.				Chronological Age		
	Range	Mean	S.D.	N	Range	Mean (in yrs.)	S.D. (in mos.)
Third Grade							
Fully-Graded Boys	74-130	104.28	14.33	53	7-11 to 9-7	8.47	4.49
Multi-Graded Boys	75-132	104.30	14.49		7-10 to 9-7	8.49	4.83
Fully-Graded Girls	81-141	107.08	14.22	51	7-10 to 9-0	8.29	3.20
Multi-Graded Girls	80-142	107.10	14.28		7-10 to 8-10	8.28	3.10
Fifth Grade							
Fully-Graded Boys	72-142	100.56	13.74	62	9-11 to 11-9	10.40	4.70
Multi-Graded Boys	70-143	100.40	13.87		9-10 to 11-9	10.38	4.91
Fully-Graded Girls	78-134	106.33	13.79	46	9-10 to 11-9	10.40	4.00
Multi-Graded Girls	77-134	105.93	13.66		9-9 to 11-10	10.42	4.36

TABLE II

RESULTS OBTAINED ON 53 PAIRED BOYS AND 51 PAIRED GIRLS AT THIRD-GRADE LEVEL; 62 PAIRED BOYS AND 46 PAIRED GIRLS AT FIFTH-GRADE LEVEL

THIRD GRADE	BOYS			GIRLS			TOTAL		
	Mean			Mean			Mean		
Subtest	Fully-Graded	Multi-Graded	t	Fully-Graded	Multi-Graded	t	Fully-Graded	Multi-Graded	t
Reading Vocabulary	3.04	2.86	1.02	3.60	3.42	1.07	3.31	3.13	1.49
Reading Comprehension	2.89	3.03	-.93	3.67	3.48	1.17	3.27	3.25	.18
Arithmetic Reasoning	2.78	2.92	-1.09	3.39	3.27	.86	3.08	3.09	-.11
Arithmetic Fundamentals	3.02	3.19	-1.79	3.13	3.26	-1.78	3.08	3.23	-2.50*
Mechanics of English	2.78	2.59	1.41	3.24	3.12	.80	3.00	2.85	1.50
Spelling	2.71	2.66	.40	3.35	3.16	1.73	3.02	2.91	.89
Battery	2.88	2.91	-.29	3.40	3.29	.90	3.13	3.10	.38
FIFTH GRADE									
Subtest									
Reading Vocabulary	5.08	4.83	1.18	6.09	5.85	1.06	5.51	5.27	1.56
Reading Comprehension	4.88	4.67	1.04	5.76	5.63	.65	5.25	5.08	1.19
Arithmetic Reasoning	5.12	5.13	-.05	5.40	5.42	-.10	5.24	5.25	-.07
Arithmetic Fundamentals	5.16	4.86	2.29**	5.18	5.19	-.08	5.17	5.00	1.79
Mechanics of English	4.60	4.74	-.78	5.91	5.98	-.28	5.16	5.27	-.75
Spelling	4.71	4.57	.61	5.74	5.95	-.80	5.15	5.16	-.06
Battery	4.93	4.81	.83	5.69	5.68	.07	5.25	5.18	.67

* Significant at .02

** Significant at .05

Discussion

The results of this study support the stated hypothesis. There is no difference in the achievement of rural children, whether they are educated in a single-graded or multi-graded school unit. Arithmetic fundamentals was the only subtest area which showed a statistically significant difference between the two groups. However, this difference was not consistent so that it is not possible to make any real generalizations in this area.

The findings in this study are in line with those of Drier (3), Adams (1), and McIntosh and Schrammel (7), while in conflict with those of Martens (6), Foote (4), and Clem and Hovey (2). Differences in design and statistical procedures and measures of achievement used may account for the differences in the findings.

One of the basic problems in a study such as this one lies in the definition of level of achievement. It is obvious that any achievement test can only measure a part of a child's educational achievements. Thus, any generalizations made on the data have to be in terms of the test instrument used as the measure of achievement.

Another problem is in the definition of a large district. By some criteria none of the schools in this study could be classified as large districts. However, in a predominantly rural community, the difference is often related to having enough teachers so that each may teach a single-grade level, rather than the multiple grades taught in schools of four teachers or less. The results may well have been different if a large district were defined as one having 5,000 or more average daily attendance.

One of the uncontrolled variables in the present study is the competency of the teacher. It is a fact that on the average, less adequate salaries are paid in districts having one, two, three or four teachers. One would expect a positive relationship between teacher competency and salary, although there may be exceptions. On the other hand, one is also likely to find a lower pupil-teacher ratio in these schools, and hence the child may be more likely to receive greater amounts of individualized instruction. These two factors may compensate for one another.

In this study, no consistent difference in achievement between multi-graded and single-graded rural children was found. On the basis of these findings, one would not expect to find differences in achievement for similar populations utilizing the same measure of achievement.

Summary

A comparison of the achievement of multi-graded and single-graded rural school children was made. Matched pairs of 53 boys and 51 girls at the third-grade level and 62 boys and 46 girls at the fifth-grade level comprised the samples studied. Arithmetic Fundamentals was the only subtest area which showed a statistically significant difference. This difference was not consistent. The findings of this study support the stated hypothesis that there are no differences in the achievement of rural school children, whether they are educated in a single-graded or multi-graded school environment.

REFERENCES

1. Adams, Joseph J. "Achievement and Social Adjustment of Pupils in Combination Classes Enrolling Pupils of More Than One Grade," Journal of Educational Research, XLVII (October 1953), pp. 151-55.
2. Clem, Orlie M., and Hovey, Chester W. "Comparative School Achievement of Village School Pupils and Rural School Pupils," Elementary School Journal, XXXIV (December 1933), pp. 269-72.
3. Drier, William H. "Differential Achievement of Rural Graded and Ungraded School Pupils," Journal of Educational Research, XLIII (November 1949), pp. 175-86.
4. Foote, John M. "A Comparative Study of Instruction in Consolidated and One Teacher Schools," Journal of Rural Education, II (April 1923), pp. 337-51.
5. Hull, J. H. "Multi-Grade Teaching," Nation's Schools, LXII (July 1958), pp. 33-6.
6. Martens, Clarence C. "Educational Achievements of Eighth Grade Pupils in One-Room Rural and Graded Town Schools," Elementary School Journal, LIV (May 1954), pp. 523-25.
7. McIntosh, H. W., and Schrammel, H. E. "Comparison of the Achievement of the Eighth Grade Pupil Both in Rural Schools and Graded School," Elementary School Journal, XXX (December 1930), pp. 301-06.
8. Nelson, Thomas L. A Comparison of the Achievement of Pupils in Schools of One or More Teachers With That of Pupils in Schools of Eight or More Teachers, unpublished doctoral dissertation (University of California, 1932).
9. Wilson, W. K., and Ashbaugh, E. J. "Achievement and Rural and Consolidated Schools," Educational Research Bulletin, VIII (1929), pp. 358-63.

personality traits. In choosing the personality tests, an attempt was made to select inventories purporting to measure so-called normal traits of personality, particularly with respect to socialization or interpersonal behavior, since prior study (1) indicates that this aspect of personality is probably related to women's attitudes toward mathematics.

In addition to the tests listed above, the Scholastic Aptitude Test (SAT) scores of each student were on file at the college, and these were used as measures of intellective ability. It should be noted that the participating students took the SAT about 18 months prior to this investigation.

Results

The product-moment intercorrelations of scores on all 42 personality variables and the Math Attitude Scale were computed.[1] The 16 statistically significant correlations of Math Attitude Scale scores and personality variables, as well as the intercorrelations among these variables, are presented in Table II. Grouping the intercorrelations of the personality variables into clusters by the ramifying linkage method (Cattell, 3), in which a minimum correlation of .50 was set for right of entry into a cluster, produced the following results. Variables 1, 2, 3, 7, 9, 10 constitute a cluster which may be labeled "extroversion," and variables 4, 5, 6, 7, 8, 10 form another cluster which will be called "conscientiousness." The theme of intellectual efficiency (items 7, 9, 10) runs through both of these clusters. Variables 5, 7, 11, 14 compose a cluster named "self-control," while variables 12 (16 PFQ Ego Strength), 13 (16 PFQ Conservatism-Radicalism), 15 (SV Theoretical Scale) and 16 (SAT-Mathematical) have no correlations above .50 with the other personality variables. Although none of the correlations of the personality variables with Math Attitude Scale scores is extremely high, only two or three of the obtained 16 correlations would be expected, by chance, to be significant at the .05 level and none at the .01 level. Thus, the results are too numerous and too internally consistent to be spurious. Generalizing, these results suggest that women with more favorable attitudes toward mathematics tend to be more outgoing, conscientious, self-controlled, intellectually mature, and place more value on theoretical matters than those with less favorable attitudes toward mathematics.

In order to determine whether these personality variables are merely related to Math Attitude Scale scores because of the common correlations of the former and the latter with mathematical ability (SAT-M), the first-order partial correlations of Math Attitude Scale scores with each of the personality variables, partialling out the effects of SAT-Mathematical scores, were computed. The results appear in Table III. It will be observed that six of the 15 correlations remain statistically significant, four at the .01 level. These results suggest that, with mathematical ability partialled out, students making higher scores on the Math Attitude Scale tend to be more socially and intellectually mature, more self-controlled, and place more value on theoretical matters than low scorers on the scale.

Summary and Conclusions

The hypothesis of this investigation was that attitudes toward mathematics are significantly related to general personality variables. In order to test the hypothesis, three objective tests of personality, the California Psychological Inventory, the Sixteen Factor Personality Questionnaire, and the Study of Values were administered to 160 college sophomore women in a general psychology class. Scholastic Aptitude Test scores were used as measures of intellective ability. The scores on these 42 variables and scores on a scale of attitudes toward mathematics with each of the personality variables significantly correlated with the former, partialling out the effects of mathematical ability, were computed. The results suggest that high scorers on the attitude scale, with mathematical ability statistically controlled, tend to be more socially and intellectually mature, more self-controlled, and place more value on theoretical matters than low scorers on the scale. These findings suggest that attitude toward mathematics is related to a broad constellation of personality variables indicative of adjustment and interest.

REFERENCES

1. Aiken, L. R., and Dreger, R. M. "The Effect of Attitudes on Performance in Mathematics," *Journal of Educational Psychology*, LII (1961), pp. 19-24.
2. Allport, G. W., Vernon, P. E., and Lindzey, G. *Study of Values, Manual* (Boston: Houghton Mifflin, 1960).
3. Cattell, R. B. "A Note on Correlation Clusters and Cluster Search Methods," *Psychometrika*, IX (1944), pp. 269-284.
4. Cattell, R. B., Saunders, D. R., and Stice, G. *Sixteen Personality Factor Questionnaire, Handbook* (Champaign, Illinois: Institute for Personality and Ability Testing, 1957).
5. Dreger, R. M., and Aiken, L. R. "The Identification of Number Anxiety in a College Population," *Journal of Educational Psychology*, XLVIII (1957), pp. 344-51.

---

[1] The facilities of the Univac 1105 at the University of North Carolina were made available for these computations.

TABLE III

SIGNIFICANT FIRST-ORDER PARTIAL CORRELATIONS OF MATH ATTITUDE SCALE SCORES WITH 15 PERSONALITY VARIABLES, PARTIALLING OUT SAT-M[a] - MATHEMATICAL SCORES

Variable	Correlation Coefficient
CPI[b] Dominance	.16*
CPI Self-Control	.22**
CPI Achievement via Conformance	.21**
CPI Intellectual Efficiency	.19*
16PFQ[c] Integration	.25**
SV[d] Theoretical Scale	.29**

[a] Scholastic Aptitude Test - Mathematical
[b] California Psychological Inventory
[c] Sixteen Personality Factor Questionnaire
[d] Study of Values
* Significant at the .05 level
** Significant at the .01 level

6. Gough, H.G. *California Psychological Inventory, Manual* (Palo Alto, California: Consulting Psychologists Press, Inc., 1956).
7. Gough, M.F. "Mathemaphobia: Causes and Treatments," *Clearing House*, XXVIII (1954), pp. 290-294.
8. Malone, W.H., and Fred, E.L. "A Preliminary Study of the Group Attitudes of Junior and Senior High School Pupils Toward Mathematics," *Journal of Educational Research*, XLVII (1954), pp. 599-608.
9. Plank, E.N. "Observations on Attitudes of Young Children Toward Mathematics," *Mathematics Teacher*, XLIII (1950), pp. 252-63.
10. Poffenberger, T.N., and Norton, D.A. "Factors Determining Attitudes Toward Arithmetic and Mathematics," *Arithmetic Teacher*, III (1956), pp. 113-16.
11. Rochlin, I. *Nonintellectual Factors in the Learning of General Mathematics*, unpublished doctoral dissertation, University of Chicago (1952).
12. Tulock, M.K. "Emotional Blocks in Mathematics," *Mathematics Teacher*, L (1957), pp. 572-76.
13. Weaver, H.B. "The Mathematical Proficiency of Students of Psychology," *Journal of Educational Research*, LI (1957), pp. 161-70.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 9, May - June 1963)

An Evaluation of the California Achievement Test, Elementary, Form W, Reading Vocabulary*

CHARLES SHEPPARD
WILLIAM J. CAMPBELL
Hicksville School District, New York

THE PURPOSE of this study is to measure the relative effectiveness of the California Achievement Test Reading Vocabulary (CAT).

The CAT may be considered as representative of the type of test available for use in school districts across the nation. This test is provided with national norms which the authors suggest should be used as reference points not as "rigid standards". It is offered in the manual that, "when the obtained scores of a testing program of a particular school or community depart significantly from the test norms, it means simply that these scores are (above or below) the average scores of a population-at-large in which the quality measured closely approximates a normal distribution curve" (13).

The problem at hand is how do the Hicksville students compare with the normative group? A frequency distribution of difficulty indices is presented for the reading vocabulary test, grades four through six. These data have not been adjusted for guessing effects. The authors state that, "There are relatively infrequent instances when guessing does have a noticeable effect on the test results of the individual" (13).

In view of the above, no attempt will be made to statistically evaluate the effect of chance. Comparison will be effected for the norm and Hicksville sample upon the difficulty indices frequency distribution presented in the manual.

Method

The CAT, Form W, elementary, was administered to 872 members of the sixth grade, 964 fifth-grade and 1070 fourth-grade students, as part of the testing procedure of the school district. These figures include the population of these grades, except those students who were absent due to illnesses.

The tests were machine scored and tabulated on the international test scoring machine graphic item counter forms. An item analysis of each item was computed. Difficulty indices for each item were determined for each grade level and track level within grade. These were organized in a frequency distribution.

Since the data provided in the test manual is for four forms of the CAT, it was assumed that each of these forms contributed equally to the difficulty indices presented by the authors. Hence, by dividing the number of forms into each total item difficulty, a distribution of item difficulty frequencies is obtained for one form.

To obtain a frequency distribution of item difficulties representative of grades four, five, and six of the Hicksville population (H), difficulty indices were totaled by interval and by grade level to ascertain a grand total of item difficulty frequencies for a specific interval. These grand totals then were divided by the number of grades (3). A frequency distribution of item difficulties for grades four, five, and six was attained for the (H) population.

Chi-square was applied comparing the frequency distributions of the CAT sample with the (H) population. Chi-square was applied to the frequency distribution of grade four of the (H) population and that of the CAT sample of grades four, five, and six.

Results

Chi-square was computed to determine if the distribution of item difficulty based on the (H)

---

* This work was in part supported by a grant from the New York State Education Department.

students data could have been drawn from a population of that of the CAT-RV norm data. A chi-square of 185.4 was attained which is significant well beyond the .001 level, with nine degrees of freedom. Therefore, the distribution of the item difficulties was not drawn from a distribution similar to that of the CAT-RV norms.

The distribution of fourth-grade students was then tested to determine if this distribution of item difficulties was drawn from that of the normative group. A chi-square of 45.5 was attained. This is significant of the .001 level. Both the combined distributions of grades four, five, and six and that of the fourth grade are significantly different from the norm data. Therefore, the distributions of the sixth and fifth are also different from the norm data. This is reflected in Table I.

Discussion

It has long been realized that there are sectional differences in performance of students on standardized tests. Because of this, the national norms offered are not always adequate as reference points for individual schools or school districts. It would be of benefit for school districts to evaluate the performance of their students on these tests. As has been shown here, there is a marked difference between performance of Hicksville students and that of the normative group.

The item difficulty distribution shown in Table I for population (H) shows a negative skew. That of the California appears mesokurtic. All distributions are significantly different (.001) from the norm distribution a composite of fourth-, fifth-, and sixth-grade students. The data compiled for fourth-grade (H) students approximate a bimodal distribution. Twenty-five items (half of the test) have a difficulty index above seventy. This, in effect, means the greater preponderance of the items is not difficult for these fourth-grade students. The power of this test is greatly minimized. A cursory examination of the distributions for the fifth- and sixth-grade students reveals a further reduction in test power at higher grade levels, as is shown in Table I.

The nature of these distributions reveals little variance. Items of median difficulty will produce greatest variance. Most of these items are closer to the "variance vanishing point". Variance is necessary to identify individual differences among the students. Because of this "lack of variance", little individual differences can be identified. Secondly, average intercorrelation of vocabulary items, as well as their discrimination indices, tend to be high. For this reason, items should be chosen over a range of ability for maximum discrimination. Again, the test fails to meet this consideration.

The argument may be advanced that this is an achievement test designed to measure the effectiveness of curriculum retention. Performance indicates the degree to which the students have mastered the curriculum. It would seem to indicate that the fourth-grade students have mastered the fifth- and sixth-grade curriculum without exposure.

The answer seems to be that the test will not do for samples of the composition evaluated here. These results can only be generalized to performance of students from school districts alike or similar to the Hicksville population.

The development of local district norms will aid in evaluating the performance of local students. A serious error is made if it is felt that this is the complete answer to the situation. This is only a "stop gap" method until a more powerful, valid test is found.

Summary and Conclusion

A comparison of performance of 2906 students was conducted with that of the norm group on the California Achievement Test, elementary, reading vocabulary. The evaluation was effected by grade compared to the norm group of a composite of fourth-, fifth-, and sixth-grade students. In all cases the distribution of difficulty indices was statistically significant from the norm distribution beyond the .001 level. This section of the test is not sufficiently powerful for this sample of students. Further data of difficulty indices are available by grade and track from the authors.

BIBLIOGRAPHY

1. Adams, J. F. "Test Item Difficulty and the Reliability of Item Analysis Methods." Journal of Psychology, XXXXIX (1960), pp. 255-262.
2. Adams, J. F. "The Effect of Non-Normally Distributed Criterion Scores on Item Analysis Techniques." Educational and Psychological Measurement, XX (1960), pp. 317-320.
3. Anastasi, A. Psychological Testing (New York: MacMillan Company, 1954).
4. Boneau, C. A. "The Effects of Violations of Assumptions Underlying the T Test." Psychological Bulletin, LVII, pp. 1, 49-64.
5. Brogden, H. E. "The Effects of Bias Due to Difficulty Factors in Product Moment Item Intercorrelations on the Accuracy of Estimation of Reliability by the Kuder-Richardson Formula Number 20. Educational and Psychological Measurement, VI (1946), pp. 517-520.
6. Cook, D. L. "A Note on Relevance Categories and Item Statistics." Educational and Psychological Measurement, XX (1960), pp. 321-

331.

7. Davis, F. N. "Item Analysis Data: Their Computation, Interpretation, and Use in Test Construction." Harvard Educational Papers, No. 22 (1946).
8. Flanagan, J. C. "General Considerations in the Selection of Test Items and a Short Method of Estimating the Product-Moment Coefficient from Data at the Tails of the Distribution." Journal of Educational Psychology, XXX (1939), pp. 674-680.
9. Nunnally, J. "The Place of Statistics in Psychology." Educational and Psychological Measurement, IV (1960), pp. 641-650.
10. Rulon, P. J. "A Simplified Procedure for Determining the Reliability of a Test by Split-Halves." Harvard Educational Review, IX (1939), pp. 99-102.
11. Scott, W. A. "Measures of Test Homogeneity." Educational and Psychological Measurement, IV (1960), pp. 751-757.
12. Toops, H. A. "A Comparison, by Work-Limit and Time-Limit, of Item Analysis Indices for Practical Test Construction." Educational and Psychological Measurement, II (1960), pp. 251-266.
13. Tiegs, E. W., and Clark, W. W. California Achievement Test Complete Battery Manual, Elementary, Grades 4, 5 and 6, Forms W, X, Y, Z. California Test Bureau 1957.

TABLE I

ITEM DIFFICULTY DISTRIBUTIONS OF CALIFORNIA NORM DATA AND HICKSVILLE DATA GRADES 4, 5 and 6

Item Difficulty Intervals	California Norms Grades 4, 5 and 6	Hicksville Grades 4*	5*	6*
90 - 100	1.5	8	15	26
80 - 89	4.2	9	16	9
70 - 79	7.0	8	5	6
60 - 69	12.8	5	4	4
50 - 59	11.2	4	5	3
40 - 49	7.5	6	3	1
30 - 39	5.2	7	2	1
20 - 29	.5	2	-	-
10 - 19	-	1	-	-
0 - 9	-	-	-	-
Total	49.9		50	

* Sixth-grade data is based on an N of 872

* Fifth-grade data is based on an N of 964

* Fourth-grade data is based on an N of 1070

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 9, May - June 1963)

A Factorial Analysis of Fields of Scholarly Production

ROBERT H. KNAPP
LINDA A. EHLINGER
Wesleyan University

IT IS A common practice among educational administrators and scholars to speak of the "natural sciences," "social sciences" and "humanities" as the major categories of scholarship. Sometimes the division between biological sciences and physical sciences is made, and in some cases there is doubt concerning the classification of a particular field such as history, which at times is grouped with the "humanities" and at other with the " social sciences". It is an administrative fact, however, that the fields of scholarly inquiry are commonly classified according to such a scheme with divisional organization and deans assigned accordingly. It is the purpose of the present article to present some empirical data from which the major groupings of scholarly fields may be derived in order to determine the soundness of the present *a priori* mode of classification.

The data upon which this study rests are derived from *Doctorate Production in United States Universities 1936-1956*, National Academy of Sciences (1). The authors of this publication have presented a list of the 30 institutions in the nation granting the greatest number of doctor's degrees during the period specified. For each institution, the number of such persons in different fields of learning is given. Our first step consisted of reducing to percentages the numbers in particular disciplines as a proportion of the total doctor's degrees granted by each institution. We took the liberty of eliminating from consideration certain disciplines such as religion, speech, business administration, and jurisprudence, since these tend to be specialized degrees concentrated in a relatively small number of our 30 graduate institutions. Thus, for each of our 30 largest graduate schools we obtained the percentage of their total doctor's degrees granted in each of 25 different fields. The resulting percentages give no information concerning the absolute number of doctor's degrees granted by different graduate schools; they do, however, give a quick picture of the fields in which the major contributions have been made, thus yielding a sort of profile which defines the main scholarly emphasis of the particular graduate institution in question.

Our next step consisted of correlating the percentage scores for each field with every other for our 30 graduate schools. Thereafter a centroid factor analysis was carried out upon the correlation matrix, and these factors were rotated according to a Varimax Program. Table I shows the loadings of each field of scholarship on the first five factors. It will be seen on inspection that the first three factors yield a large number of substantial loadings, while the last two yield only small residual loadings, for the most part. But the importance of the first three factors, from a statistical point of view, is beyond dispute.

It will be seen that the first factor, tentatively called a "physical science" factor, yields high loadings (in excess of +.60) in the fields of mathematics, physics, and earth science. At the opposite pole of this factor stands the field of education with a negative loading of -.70. Thus, our first factor reveals the tendency for high production in mathematics, physics, and earth science to be associated together in the same institution, while high production in education is negatively correlated with these.

The second factor is essentially a "biological

---

[*] This research was supported by the Council of the Humanities of Princeton University under a grant from the Ford Foundation.

TABLE I

ROTATED FACTOR LOADINGS OF 25 FIELDS OF SCHOLARSHIP

Field	Factor I	Factor II	Factor III	Factor IV	Factor V
Mathematics	+.72	-.11	+.12	-.33	+.04
Physics	+.72	-.17	-.20	+.10	-.19
Chemistry	+.41	+.17	-.63	+.06	-.06
Earth Science	+.63	-.35	-.01	00	+.23
Engineering	+.45	-.03	-.68	+.22	-.29
Agriculture	+.04	+.86	-.32	-.05	-.15
Botany	-.05	+.90	+.08	-.04	+.18
Biochemistry	-.20	+.69	-.14	+.06	+.06
Genetics	+.03	+.88	-.31	00	-.07
Microbiology	-.04	+.58	-.19	+.36	+.12
Physiology	-.03	+.38	00	+.12	+.62
Zoology	-.08	+.05	+.16	+.27	+.43
Medical Science	+.05	+.08	+.05	+.61	+.15
Sociology	-.21	+.02	+.64	-.02	-.14
Anthropology	+.08	-.12	+.45	+.59	+.19
Economics	+.08	+.51	+.47	-.11	-.09
Geography	+.01	-.05	+.03	+.25	+.61
History	-.06	-.17	+.82	-.01	-.03
Political Science	+.29	-.21	+.57	-.11	+.50
Psychology	-.42	-.20	-.10	-.29	+.29
Arts and Music	-.17	-.22	+.21	-.35	+.39
Foreign Languages	+.28	-.26	+.79	+.21	-.17
English Literature	+.16	-.23	+.72	+.18	+.23
Philosophy	-.04	-.15	+.58	-.05	-.34
Education	-.70	-.44	-.07	-.16	-.12

sciences'' factor. Here we find agriculture, botany, biochemistry, genetics, and microbiology, all with positive loadings in excess of .69. But this factor seems particularly associated with plant biology, and it is interesting to note that zoology, physiology, and medical science have but small loadings on this factor. Interestingly enough, the field of economics yields a higher loading, by far, on this factor than any of the biological fields oriented toward animal, as opposed to plant, materials. Finally, it is worth passing note that the field of education yields a moderately high negative loading on this factor, indicating that here, as on our first factor, productivity in the field of education is negatively related to scientific productivity.

Our third factor is clearly a ''humanities'' factor. It yields positive loadings in excess of .57 in six non-scientific fields, namely, history, foreign languages, English literature, sociology, philosophy, and political science, with the largest loadings on the first three of these. But it is particularly interesting to note that two of these six, namely, political science and sociology, are commonly classified as ''social science'', and, indeed, history is frequently so classified. Thus, while we call this factor a ''humanities'' factor, it should be noted that the so-called ''social sciences'' figure in the picture, while, on the other hand, arts and music do not. This factor is actually bi-polar, as is our first factor. It will be observed that while high positive loadings tend to occur in the traditional humanistic disciplines, two fields, namely, chemistry and engineering, have very substantial negative loadings. Thus, we find opposed and negatively correlated here the humanities and the two fields most associated with physical technology. It is interesting to see that these two fields show more striking negative loadings on this factor than positive loadings on our first factor which is oriented primarily to the physical sciences. Those who distinguish sharply between physical science and technology will find confirmation of this distinction in these results.

Our fourth and fifth factors may be rather quickly dismissed since, for the most part, the loadings here are negligible or involve fields which are unusually small. In passing, however, we should observe that both medical science and anthropology show substantial positive loadings on the fourth factor, suggesting a kind of symbiosis between these fields, a proposition which is not implausible. On the fifth factor, again, we may note that physiology, geography and political science yield moderate negative loadings in excess of .50, suggesting some possible kinship of these fields. In the main, however, the last two factors are clearly not as saliently defined as the first three, and can be dismissed as merely suggestive.

Let us now summarize quickly the findings which emerge from this analysis. We have sought here to identify the factorial relations of 25 fields of scholarship in terms of the degree to which scholarly production of doctor's degrees is related. We have discovered three factors of prime importance and two other suggestive factors. Our first factor is bi-polar and shows the essential opposition of the physical sciences and the field of education. The second factor groups together biological sciences, excepting those particularly concerned with animal, as opposed to plant life, and shows again the opposition of these to the field of education. Our third factor groups together most of the traditional humanistic disciplines with two of the social sciences, sociology and political science, though these are but marginally identified with this cluster. In opposition to these humanistic fields stand the technologically oriented fields of chemistry and engineering.

Let us now return to the traditional patterning of the ''sciences'', ''social sciences'', and ''humanities''. If the data which we have employed constitute a legitimate point of departure, then it is clear that the conventional grouping is at least in some respects defective. Among the so-called ''sciences'' it is clear that a major distinction should be made between physical and biological sciences. Further, within each of these there are important distinctions. Thus, ''physical science'', should be distinguished from ''physical technology'' since the former attain their highest loading on the first factor, and the latter on the third factor. Among the biological sciences, those oriented to ''plant'' subject matter keep close company, but they are not significantly associated with physiology, zoology, and medical science, which latter groups appear to move in relatively isolated orbits of their own.

The homogeneity of the so-called ''social sciences'', it would appear from this data, is largely mythical. In the first place, we have no ''social science'' factor emerging, even after five factors are extracted from our matrix. Almost all of these fields, to be sure, have low to moderate positive loadings on the third, or ''humanities'' factor, and one of them, history, is clearly a pillar of the ''humanities'' factor itself. This trend, if anything, confirms the proposition that the social sciences are not positively, and perhaps even negatively, associated with the traditional sciences, as an examination of the loadings of these fields on the first and second factors will confirm. But the really striking thing is their heterogeneity. Thus, economics has its highest loading on the second factor, sociology and political science on the third, anthropology on the fourth, and geography on the fifth factor. Psychology shows an extraordinary pattern of independence with only a moderate tendency to be associated with the field of education on the first factor, in opposition to the physical sciences.

And now, how homogeneous are the ''humanities''?

Here traditional modes of classification are not so at variance with our findings if history be called one of the humanities. It is clear that history, foreign languages, English literature, and philosophy keep common company. But they do not include in their ranks arts and music, while they accept to tentative membership in their fraternity sociology and political science. Despite these exceptions, it would appear that here the traditional a priori classification is fairly well vindicated.

What about the field of education? Are administrators sound in thinking that it constitutes a special department apart from the traditional fields of scholarship? It would appear so. An inspection of our table of factor loadings will show that there is nothing like education. It is negatively related to the physical sciences in a high degree, to the plant-oriented biological sciences in a lesser degree, and elsewhere claims no affinity with other fields. Its position in our factorial space is, therefore, unique, and its only identification is found in the fact that it avoids the company of both the physical and biological sciences.

One final observation should be made concerning this data. Two hypotheses have been invoked to describe the possible relation between closely allied fields of scholarship. The first is the complementary hypothesis which would argue that strength in one department is attained at the cost of development in closely allied departments. Thus, a strong mathematics department might actually impede the scholarly production of a physics department by attracting able students from them. Or again, a strong history department might weaken an English department through the same principle. A second hypothesis, which we shall call the supplementary hypothesis, is that, in the main, strength in an allied field facilitates the output of scholars in a given field. The data which we have presented strongly suggest that in the main the latter principle prevails. Otherwise we should never have emerged with the factorial picture we have presented. Thus, to strengthen one of the physical sciences, biological sciences, or the humanities, seems likely to strengthen disciplines closely allied to them as well, rather than to work to their disadvantage through a competitive or complementary principle.

REFERENCE

1. National Academy of Sciences - National Research Council. Doctorate Production in United States Universities 1936-1956 (Washington, D.C.: Publication No. 582).

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 9, May - June 1963)

The Relation of Mental Health Analysis Scores to Teacher Ratings of Mental Health Status of Elementary Pupils

JOSEPH C. BLEDSOE
University of Georgia

THIS STUDY is concerned with whether elementary pupils selected by their teachers as having relatively high mental health status and relatively low mental health status differ with respect to the several components, categories, and total score of the Mental Health Analysis. The correlation of teacher ratings of mental health status to the respective parts and total score of the Mental Health Analysis was also studied.

Procedures

The first task involved the development of a technique for the identification of the pupils at the extremes of the "mental health" continuum. In view of the variability of the expression "mental health," it seemed fitting to employ a procedure based upon extended observation of the pupils involved. Nomination by teachers was selected as the most feasible and appropriate procedure. The criteria of mental health used by the teachers when making the nominations were developed in a preliminary study. For more than a year prior to the present study, the writer was engaged with the faculty of an elementary school in an action-research program designed to develop greater insight and understanding of mental health as it is related to the school. As a part of the study, weekly staff conferences were held to develop common definitions, some guiding principles, and a consensus concerning the role of the teacher in promoting the mental health of his pupils. For purposes of the study, mental health was defined as "a relatively enduring state wherein the person is well-adjusted, has a zest for living, and is attaining self-actualization or self-realization." (3)

Mental Health Analysis

This questionnaire, prepared by Louis P. Thorpe and Willis W. Clark and published by the California Test Bureau, consists of 200 items intended to assay the degree of mental health possessed by the examinee. Separate booklets are available for four levels of maturity — elementary (grades four to eight), intermediate (grades seven to ten), secondary school and college (grades nine to sixteen), and adult. Each of the 200 questions at each level is to be answered by circling Yes or No. Five sorts of personality "liabilities" are investigated by 100 of the questions. These liabilities are described as (a) behavioral immaturity, (b) emotional instability, (c) feelings of inadequacy, (d) physical defects, and (e) nervous manifestations. The other 100 questions are designed to estimate the examinee's personality "assets," which are designated as (a) close personal relationships, (b) interpersonal skills, (c) social participation, (d) satisfying work and recreation, and (e) adequate outlook and goals. The items are so classified that 20 apply to each specific classification.

Reliability coefficients vary from .87 for the Secondary Liabilities Score to .95 for the Adult Total Score. For the Elementary Scale which was employed in this study, the reliability was reported as follows: Assets, .90; Liabilities, .89; Total, .90. Reliability coefficients for component scores of the Elementary Scale vary from .80 for "outlook and goals" to .85 for "physical defects." (2)

Content validity of the instrument is based upon discussion of the adequacy of item selection, the meaningfulness of their analysis of the mental health categories, and the cleverness in disguise of items. Concurrent validity is reported for studies of mental health factors related to social acceptance or social rejection of elementary pupils (1), delinquency or non-delinquency of junior-high boys (8), sociomoral behavior of high school boys (7), and efficiency of psychiatric aides in hospitals (6). These studies gave some indication of the

TABLE I

COMPARISON OF PUPILS RATED BY TEACHERS AS HAVING HIGH MENTAL HEALTH STATUS WITH THOSE RATED AS HAVING LOW MENTAL HEALTH STATUS ON COMPONENTS AND TOTAL SCORES OF THE MENTAL HEALTH ANALYSIS

Mental Health Analysis Component	Teacher-Rated High Mental Health Status (N=34)		Teacher-Rated Low Mental Health Status (N=30)		Diff.	t
	$\overline{X}$	S.D.	$\overline{X}$	S.D.		
Assets	87.1	11.94	72.7	10.81	14.4	7.13**
Close Personal Relationships	18.0	2.08	14.5	2.40	3.5	3.90**
Inter-Personal Skills	17.9	2.69	13.9	2.40	4.0	4.27**
Social Participation	17.8	2.41	14.6	1.73	3.2	3.71**
Satisfying Work and Recreation	17.1	2.76	15.2	3.26	1.9	1.83
Outlook and Goals	16.3	3.32	14.5	2.90	1.8	1.70
Liabilities	80.8	9.45	67.8	12.45	13.0	6.90**
Behavioral Immaturity	16.8	2.63	14.4	3.11	2.4	2.36*
Emotional Instability	14.3	2.31	11.3	3.06	3.0	3.05**
Feelings of Inadequacy	14.1	1.99	11.2	2.90	2.9	3.08**
Physical Defects	19.3	2.64	16.7	2.69	2.6	2.66*
Nervous Manifestations	16.3	1.98	14.2	2.89	2.1	2.24*
Total Mental Health Analysis	167.9	20.1	140.5	21.1	27.4	10.06**

* Significant at .05 level
** Significant at .01 level

ability of the Liabilities category and components, the Assets category, and the total score to differentiate between groups of individuals judged to differ in adjustment patterns by an outside criterion.

The Subjects

Prior to the administration of the Mental Health Analysis inventory, teachers were asked to list the three (approximately) pupils who were considered to have the highest health status and the three (approximately) who were deemed to have the lowest health status (as defined in the study). The teachers involved were asked to name those who "stood out" in their minds without restricting to an exact number, since the intent was to define two groups representing the extremes of the distribution. The teachers submitted these lists at a point in the school year when they had observed the children for about four months. Elementary Form A of the Mental Health Analysis was then administered to all pupils.

A total of 34 pupils was rated by their teachers as having "high" mental health status and 30 pupils were rated as having "low" mental health status. In addition to these groups, there were 133 pupils in the "middle" group, making a total of 197 subjects within the three categories.

Statistical Treatment

Means and standard deviations for the two extreme groups were computed on the Mental Health Analysis. Differences were then tested for significance by means of the t-test of significance. The .05 level of confidence was accepted as the criterion for rejection of the null hypotheses involved. The correlations of teacher ratings to pupil performance on the Mental Health Analysis (referred to hereinafter as pupil self-ratings), were determined by use of the coefficient of contingency resulting from a three-by-three chi-square contingency table.

Results

Table I presents the comparison of scores on the Mental Health Analysis for the groups rated by teachers as "high" and "low" in terms of mental health status. All differences favor the group rated "high." Eleven of the 13 differences exceed the .05 level of significance. Within the limitations of this study, these findings suggest that teachers can, following a period of in-service education, differentiate between pupils of varying mental health status. If one would assume the validity of teacher ratings of mental health, the pupils' responses on the Mental Health Analysis would appear to be validated by these teacher ratings.

Table II presents contingency coefficients for the teacher-rated and pupil self-rated mental health status for the total group of 197 students. The three-by-three chi-square contingency table included the "high" (upper ten percent), "middle" (middle 80 percent), and "low" (lowest ten percent) on the two sets of ratings. Eight of the 13 coefficients were significant at the .05 level or beyond. Highest relationships were in the "Feelings of Inadequacy" subscale and the MHA Liabilities scale.

Summary

This investigation is concerned with the relationship of elementary pupils' responses to the Mental Health Analysis to the categorization of "mental health status" by the teacher. Means on the areas of the Mental Health Analysis for 34 pupils ranked by teachers as being in the lowest ten percent in mental health status were compared with means of 30 pupils ranked in the highest ten percent. All means favored the group adjudged as having highest status and 11 of the 13 differences exceeded the .05 confidence level. Coefficients of contingency for the three-by-three classifications of the 197 pupil self-ratings and teacher ratings (lowest 10 percent, middle 80 percent, and highest 10 percent) were computed. Eight of 13 coefficients were significant at the .05 level, and three were .50 or greater.

REFERENCES

1. Baron, R. B. D. Mental Health Factors Associated with Social Acceptance and Social Rejection, unpublished doctor's dissertation, University of Southern California, Los Angeles (1948).
2. California Test Bureau. Manual for Mental Health Analysis (Los Angeles, 1946). (Revised 1959)
3. English, H. B., and English, A. C. A Comprehensive Dictionary of Psychological and Psychoanalytical Terms, revised (New York: Longmans, Green, 1948).
4. Jahoda, M. Current Concepts of Positive Mental Health (New York: Basic Books, 1958).
5. Mitchell, J. C. "A Study of Teachers' and Mental Hygienists' Ratings of Certain Behavior Problems of Children," Journal of Educational Research, XXXVI (1943), pp. 292-307.
6. Vaccaro, J. J. A Study of Psychological Factors that Contrast the Most and Least Efficient Psychiatric Aides in a Mental Hospital, unpublished doctor's dissertation, Fordham University (1951).

7. Warner, B. E. *Relationship of Some Factors in Sociomoral Behavior to Some Indices of Health Status of Senior High School Boys*, unpublished doctor's dissertation, University of Southern California, Los Angeles (1954).

8. Zakolski, F. C. "Studies in Delinquency: I. Personality Structure of Delinquent Boys," *Journal of Genetic Psychology*, LXXIV (1949), pp. 109-17.

TABLE II

COEFFICIENTS OF CONTINGENCY BETWEEN SCORES ON THE MENTAL HEALTH ANALYSIS SELF-INVENTORY AND THE TEACHER'S ESTIMATE OF MENTAL HEALTH STATUS OF 197 ELEMENTARY PUPILS

Assets	.411*
Close Personal Relationships	.336*
Inter-Personal Skills	.509**
Social Participation	.212
Satisfying Work and Recreation	.423*
Adequate Outlook and Goals	.332
Liabilities	.545**
Behavioral Immaturity	.488**
Emotional Instability	.229
Feelings of Inadequacy	.598**
Physical Defects	.224
Nervous Manifestations	.294
Total Mental Health Analysis	.423*

* Significant at .05 level

** Significant at .01 level

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 9, May - June 1963)

The Effects of Chronological Age in Months on School Achievement

DONALD J. DICKINSON
Oklahoma State University
J. DONALD LARSON
Sioux Falls Public School System

THE OPTIMUM AGE to enter a child in the first grade is a question that has plagued educators for years. Many educators have felt that some children encounter difficulty throughout school as a result of entering the first grade before they are mature enough to do first grade work (Brenner, 1957). In a study by Carter (1956), it was found that underage children (not six years old before entering school) did not achieve as well as the regular school age children. On the other hand, Kirk (1958) found that mentally retarded children who were given a preschool education had higher IQ scores and social maturity scores than those who entered school at a regular age. There are studies which also demonstrated that children who attend kindergarten have higher IQ scores than those who do not attend (Goodenough, 1951).

This study was performed to determine the effects of chronological age (CA) at time of entering school on later school achievement. The children were all of normal age for entering the first grade when they entered school. This differs from Carter's (1956) study in that he used children who were underage and compared them to the normal school age child.

The subjects were 480 fourth grade students attending the Sioux Falls Public School System. These children were selected by stratified random sampling from a larger population in order to control for such variables as kindergarten experience, and experience in another school system. There were no repeaters in the study. The composite score of the Iowa Test of Basic Skills (ITBS) was used to measure achievement.

Two approaches were used to determine the influence of CA on achievement. The first approach was to compare the younger fourth of the class to the remainder of the class. The second approach was to divide the class into four groups on the basis of age, and then compare the younger fourth to each of the remaining three groups (of fourths). It was hypothesized that those children who were younger in CA would differ in achievement from the older children.

For the first part of the study the subjects were assigned to the following groups according to age in months:

	Age in Months
Group I	107 to 109
Group II	110 to 118

Thus Group I consisted of children who had the youngest CA, and who entered the first grade before they were six years old or who became six years old before the 1st of November. Group II was composed of children who were at least six years in CA when they entered the first grade.

An analysis of variance supported the hypothesis that CA affects school achievement as measured by the ITBS. It was found that the younger fourth of the class had a significantly lower mean composite score of 4.73 on the ITBS than did the remainder of the class, with a mean composite score of 4.92 ($p < .05$). IQ could not account for the older group having a larger composite score, as the younger age group had a higher mean IQ than the older group, although it was not significant at the .05 level. Mean mental age, on the other hand, was found to be significantly higher for the older age group. The older group had a mean mental age of 127.03 months compared to a mean mental age of 122.34 months for the younger group ($p < .01$).

The older children's scoring higher on the ITBS could be simply a result of having a higher mental age. However, the fact that these differences still exist at the fourth grade level may point to a "snowballing" effect. The difference that existed at early ages may become magnified as the child becomes older. A child who is able to achieve at the first grade level may receive positive reinforcement for his achievement, and as a result he may be more motivated to do well in school (McDavid, 1959).

For the second part of the study, the children were divided into four groups to allow for a comparison of children who differed in a three months age range, as follows:

TABLE I

TABLE OF MEANS FOR THE THREE-MONTH-AGE-RANGE-GROUPS OF ACHIEVEMENT, MENTAL AGE, AND IQ

Groups	Composite	Mental Age	IQ
A	4.68	122.34	113.82
B	4.78	125.50	112.99
C	4.97	127.62	112.12
D	4.87	127.98	109.12

TABLE II

DISTRIBUTION OF HIGH AND LOW IQ'S IN THE THREE-MONTH-AGE GROUP

		Groups			
		A	B	C	D
IQ under 100	M	6	12	14	18
	F	7	5	6	7
Total		13	17	20	25
IQ over 130	M	7	6	5	2
	F	4	5	5	0
Total		11	11	10	2

	Age in Months
Group A	107 to 109
Group B	110 to 112
Group C	113 to 115
Group D	116 to 118

An analysis of variance revealed an F ratio of 2.45. With three and 475 df an F of 2.62 is significant at the .05 level. Although there was no significant difference between the means of the four three-month-age range groups on achievement, their differences did approach significance. An inspection of Table I reveals that achievement increased from the youngest group (Group A) through group three (Group C) and then dropped slightly on the oldest group (Group D).

When an AOV was performed using IQ, it was found that the youngest group (Group A) and next to the youngest group (Group B) had a significantly higher IQ than the oldest group (Group D) ($p < .05$).

Theoretically, all groups should have approximately the same mean IQ, as IQ is a relative measure based on CA and mental age. An examination of the data revealed that there was probably a selective process that accounted for the younger groups having higher IQ's than the older groups.

As it was shown that the younger children did not achieve as well as the older children, it appears probable that the younger children would be the most likely to fail a grade. It also seems likely that the children with the lower IQ's would achieve less than the brighter children in this younger group; and as a result, these children would be the first to fail. This leaves only the brighter children in the younger group, resulting in higher IQ's as a group.

When the data was inspected, it was found that there were thirteen children in the youngest group with an IQ of less than 100. This compares to twenty-five in the oldest three-month-age-range. (See Table II). When a chi square was applied, it was found that this difference was significant at the .01 level ($\times^2 = 7.58$).

It was also found that as the groups increased in CA, there was an increase in the number of children who had an IQ below 100. Thus it appears that if a child has an IQ less than 100, his chances of not failing a grade are better the older he is when he enters school. This is shown in Table II. It is interesting to note that there is little difference between the four groups on the number of children whose IQ's exceed 130, except on the oldest group (Group D), where there is a drop. There is the possibility that the lack of children in the older group (Group C) with IQ's in excess of 130 could be a result of those children with high IQ's and high achievement receiving a double promotion.

In spite of the younger children having higher IQ's than the older children, they are still not able to achieve as well as the older children with lower IQ's. This could be a result of the CA of the children when they enter school. To put it another way, it could be a result of the mental age of the children when they enter school as mental age should increase in proportion to CA. However, an inspection of Table I shows that mental age does not increase proportionally to achievement on the composite score of the ITBS. There is only a slight increase in achievement between groups A and B, yet it is between these two groups that the largest difference in mental age appears.

It appears that mental age may be a much better predictor of achievement than IQ at the fourth grade level. As mental age increases, so does achievement on the ITBS. IQ, on the other hand, tends to decrease. It was noted that the females scored significantly higher than males on achievement, IQ, and mental age ($p < .05$). There were, however, fewer girls than boys with an IQ of 130 or higher, and there were more boys than girls with IQ's less than 100.

REFERENCES

1. Brenner, Anton. "Nature and Meaning of Readiness for School", Merrill-Palmer Quarterly, Spring 1957, 114-35.
2. Carter, Lowell B. "The Effects of Early School Entrance on the Scholastic Achievement of Elementary School Children in the Austin Public Schools," Journal of Educational Research, L, (1956), pp. 91-1-3.
3. Goodenough, Florence. "A Critique of Experiments on Raising the IQ", Readings in Child Psychology, Prentice-Hall, Inc., (1951), pp. 320-28.
4. Kirk, Samuel A. An Early Education of the Mentally Retarded, Illinois Press, 1958.
5. McDavid, John Jr. "Some Relationships Between Social Reinforcement and Scholastic Achievement," Journal of Consulting Psychology, XXIII, (1959) pp. 151-54.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 9, May - June 1963)

Intelligence and Gain in Reading as Related to Gains in the Sub-tests of the Stanford Achievement Test

CARRIE M. SCOTT
Bend, Oregon

THIS STUDY extends a former investigation (3) in which the writer correlated intelligence scores with gains in reading achievement, and each, in turn, with gains in arithmetic reasoning, social studies and science, using 670 fifth- and sixth-grade pupils of the Bend Public Schools, Bend, Oregon, over a 19-month period.

The purposes of this study were to learn 1) whether a relationship exists between intelligence scores and gains in the sub-tests of the Stanford Achievement Test Battery, (2) namely: Paragraph Meaning, Word Meaning, Spelling, Language, Arithmetic Reasoning, Arithmetic Computation, Social Studies, Science, and Study Skills: 2) the extent to which gains in the Paragraph Meaning test and gains in the Word Meaning test are related to gains in the other sub-tests of the battery; and 3)whether a higher correlation exists if the results of the advanced battery, when administered to the pupils at the end of the sixth grade, are used in the place of those of the intermediate battery used in the previous study.

Procedure

To accomplish these ends, 193 pupils in the fifth and sixth grades of the Bend Public Schools, were the subjects for all of whom intelligence scores were available, and who had taken the Stanford Achievement Test Intermediate Battery, Form J, at the beginning of grade five in the fall of 1959 and the advanced battery, Form K, at the end of grade six in the spring of 1961.

The intelligence test scores were selected in the order of preference of Stanford-Binet Intelligence Scale (5), WISC (7), and California Test of Mental Maturity (4). The intelligence test scores together with the grade scores on the nine sub-tests of the Stanford Achievement Test battery according to both administrations were recorded. The grade scores were converted into K-scores to insure greater accuracy. Differences in K-scores were found for each pupil over the 19-month period.

The pupils were in two elementary schools and had a total of nine regular fifth-grade teachers and ten sixth-grade classroom teachers. The national norms at the times of testing were grade 5.2 and 6.8. The average obtained score for the fifth grade was 5.4 and for the sixth grade, 7.0. The fifth grade averaged 5.7 in each of the reading tests and 4.6 in arithmetic computation. The sixth grade average was highest in science, followed by paragraph meaning. The mean intelligence score for the group was 105. Gains in achievement on the sub-tests are shown in Table I.

The greatest amount of gain was made on the arithmetic sub-tests with science following. The least amount of gain was made in the language test, with word meaning and spelling next lowest. The greatest amount of variation was in arithmetic computation, and the least variability was in spelling.

Using the Pearson Product-Moment formula (1), correlations were obtained between intelligence scores and gains in the sub-tests, between gains in paragraph meaning and the remaining sub-tests, and between gains in word meaning and the remaining sub-tests. The results appear in Table II.

All of the correlations were low. The highest correlation was obtained between intelligence and gain in paragraph meaning followed by intelligence and gain in arithmetic computation. The correlations between intelligence and gains in science and study skills were negative. Gain in paragraph meaning correlated highest with gain in word meaning and with the arithmetic sub-tests. Gain in the word-meaning test correlated highest with gain in arithmetic computation. The correlations with gains in the language test all were negative.

In order to determine the relationship existing between gains in the reading tests and gains in the other sub-tests, partial correlations were computed with the effect of the intelligence scores held constant. The results of the partial correlations appear in Table III.

The partial correlations show that both intelligence and gain in reading proficiency are related to gains in the other curriculum areas as measured by the achievement test.

Discussion

A recent research study by the writer and another by Tilton (6) have revealed that correlations obtained between gains on the sub-tests of an achievement test battery are low. Other investigators have obtained higher correlations between sub-tests of a battery by using the actual test scores. Tilton observed that the brighter pupils seemed to mark time

TABLE I

MEAN GAINS IN ACHIEVEMENT ON THE SUB-TESTS OF THE STANFORD ACHIEVEMENT TEST BATTERY FOR 193 PUPILS OVER A 19-MONTH PERIOD ACCORDING TO K-SCORES TOGETHER WITH THEIR STANDARD DEVIATIONS

Areas	Means	Standard Deviations
Paragraph Meaning	9.70	7.65
Word Meaning	8.65	7.47
Spelling	8.27	4.29
Language	4.95	5.91
Arithmetic Reasoning	12.79	7.06
Arithmetic Computation	29.80	12.80
Social Studies	10.45	7.53
Science	12.48	8.88
Study Skills	9.76	8.87

TABLE II

CORRELATIONS BETWEEN INTELLIGENCE SCORES AND GAINS IN THE SUB-TESTS OF THE STANFORD ACHIEVEMENT TEST AND BETWEEN GAINS IN THE PARAGRAPH MEANING TEST AND GAINS IN THE WORD MEANING TEST WITH THE REMAINING SUB-TESTS OF THE BATTERY TOGETHER WITH THEIR PROBABLE ERRORS

Areas	Correlations	Probable Errors + or -
1. Intelligence and		
Paragraph Meaning	.241	.046
Word Meaning	.179	.045
Spelling	-.080	.048
Language	-.214	.146
Arithmetic Reasoning	.125	.048
Arithmetic Computation	.298	.044
Social Studies	.050	.048
Science	.084	.048
Study Skills	-.067	.048
2. Gain in Paragraph Meaning and		
Word Meaning	.340	.042
Spelling	.081	.048
Language	-.286	.044
Arithmetic Reasoning	.269	.045
Arithmetic Computation	.256	.045
Social Studies	.144	.048
Science	.015	.048
Study Skills	.037	.048
3. Gain in Word Meaning and		
Spelling	.067	.048
Language	-.015	.048
Arithmetic Reasoning	.143	.046
Arithmetic Computation	.204	.046
Social Studies	.159	.047
Science	.119	.047
Study Skills	.102	.048

TABLE III

PARTIAL CORRELATIONS BETWEEN GAIN IN PARAGRAPH MEANING AND WORD MEANING SUB-TESTS OF THE STANFORD ACHIEVEMENT TEST BATTERY WITH THE OTHER SUB-TESTS, HOLDING THE INTELLIGENCE QUOTIENTS CONSTANT

Areas	Correlations
Gain in Paragraph Meaning and	.311
Word Meaning	.103
Spelling	-.246
Language	.257
Arithmetic Reasoning	.200
Arithmetic Computation	.136
Social Studies	-.004
Science	-.022
Study Skills	
Gain in Word Meaning and	.096
Spelling	-.095
Language	.109
Arithmetic Reasoning	.116
Arithmetic Computation	.142
Social Studies	.097
Science	.132
Study Skills	

in the upper elementary grades while the duller ones were catching up. In the previous study as well as this investigation the writer found that many pupils of lower capacity made greater gains than did those of higher capacity. It was thought that a test with a higher "top" would raise the gains of the brighter pupils, but this made no difference. The brighter pupils rated high on the beginning fifth-grade battery and high on the battery given in the sixth grade, but did not as a rule show a great deal of gain. It made no difference whether the intermediate test was used or the advanced test which gives more room at the top.

Improvement in language, according to this test, was negatively correlated with intelligence and with gains in reading. Spelling gain was negatively correlated with intelligence scores, but positively correlated with gains in both reading tests. Gain in study skills was negatively correlated with intelligence scores, but positively correlated with gains in both reading tests.

Findings

1. Pupils were inconsistent as the amount of gain as measured by the achievement test and their intelligence scores.
2. Individual pupils were inconsistent in the amount of gains made among the different sub-tests of the battery.
3. The average intelligence rating was slightly above the national average.
4. Average grade scores on the two sets of achievement tests indicated that the pupils rated above the national norms.
5. The advanced test given at the end of the sixth grade did not yield higher correlations than the intermediate test previously administered.
6. The correlations were lowered, but still positive when intelligence scores did not complicate the investigation correlating gains in the reading tests with gains on the sub-tests.
7. Correlations between gains are necessarily lower and much less reliable than the scores upon which the gains are computed since more unmeasured variables are present.
8. Teaching aimed to provide opportunity for all to learn may cause pupils of higher ability to halt progress while those of lesser ability are making larger gains. Due to the heterogeneity found in a classroom, there probably is small value to be derived from correlating individual gains.
9. Gain in reading proficiency is related to gains in some subject areas and seems to have no relationship with other areas.
10. Research has proven that a relationship exists between intelligence and gains in the subject areas, and between gains in reading and in the other curriculum areas, but the extent of these influential factors is difficult to determine.
11. A similar investigation should be made when the next revision of the Stanford Achievement Test is issued.

REFERENCES

1. Garrett, Henry E. Statistics in Psychology and Education (New York: Longmans, Green and Co. Inc., 1958).
2. Kelley, Truman L. et al. Stanford Achievement Test (Chicago: World Book Company, 1953).
3. Scott, Carrie M. "The Relationship Between Intelligence Quotients and Gain in Reading Achievement with Arithmetic Reasoning, Social Studies and Sciences", The Journal of Educational Research, Vol. 56 (February 1963).
4. Sullivan, E. T., Clark, W. W. and Tiegs, E. W. California Test of Mental Maturity (Los Angeles: California Test Bureau, 1953).
5. Terman, Lewis M. and Merrill, Maude A. The Revised Stanford-Binet Intelligence Scale (Chicago: Houghton Mifflin Company, 1937).
6. Tilton, James Warren. "The Intercorrelations Between Measures of School Learning," The Journal of Psychology LXXXV (April 1953).
7. Wechsler, David. Wechsler Intelligence Scale for Children, (New York: The Psychological Corporation.)

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 9, May - June 1963)

Development of Responsibility Perceptions of Lower and Middle Class Children

MICHAEL ZUNICH
Southern Illinois University

INASMUCH AS there have been numerous studies concerning the development of responsibility of young children (3, 4, 5, 6, 8), comparatively little empirical evidence is found where investigators have obtained information concerning perceptions which children themselves hold toward the development of responsibility. Frequently these studies have been concerned with the measurement of parental attitudes and have neglected to measure children's perceptions of the development of responsibility.

This study was designed a) to obtain information concerning perceptions which children hold toward the development of responsibility, and b) to test the hypothesis that perceptions concerning the development of responsibility are independent of social class.

Method

The subjects consisted of 508 sixth-grade children living in Texas. Two hundred and fifty-two of the subjects were classified as lower-class (128 boys and 124 girls) and 256 as middle-class (130 boys and 126 girls) according to the McGuire-White (7) Index of Social Status (Short Form).

A Children's Responsibility Inventory, developed by Walters, Stromberg, and Lonian (9), was administered to the sixth-grade children during regular class sessions. The inventory consisted of 25 items measuring perceptions of responsibility such as care of clothing, cleanliness of self, and household chores. The children were asked to respond to such statements as : At what age do you think most boys and girls could a) take something back which you borrowed from a friend without being watched, b) put your dirty clothes in a hamper without being told, c) wash your hands before each meal without being reminded.

Walters, et al. (9) constructed the Children's Responsibility Inventory from an instrument measuring adult perceptions concerning the development of responsibility. A panel of six judges, who were members of the faculty at Oklahoma A. and M. College, determined the item validity by ratings.

In revising the instrument, Walters (10) had children at various grade levels participate to determine the age level at which their perceptions would be reliable. Children were interviewed to obtain information concerning the process by which children base their responses, and to determine misconceptions of meaning. After the interviews were conducted, it was evident that children in the sixth grade provided reliable responses.

To utilize the instrument with sixth-grade children, Walters (10) used the test-retest method to obtain a measure of reliability. When the inventories were completed one week apart, he found that the consistency of responses of 31 pupils yielded a percent of agreement of 77.

Results and Discussion

Table I shows the median ages at which lower and middle-class children responded that they believed boys and girls should be able to assume given responsibilities. With the exception of Item 20, the data failed to support the hypothesis that perceptions concerning the development of responsibility are independent of social class, suggesting that perceptions concerning the development of responsibility are dependent upon social class.

An examination of the median responses of lower and middle-class children evidences much difference in their perceptions. The responses of the lower-class children differed from those of the middle-class by more than a year (with the exception of item 20), indicating substantial disagreement. These differences are found in the following categories: a) household chores, b) children's relationships, c) care of clothing, d) assuming responsibility for the performance of activities alone, e) cleanliness of self, and f) playing alone.

Seven of the observed differences relate to household chores (Items 7, 8, 9, 10, 11, 12, 13). For example, the lower-class children indicated that eight years of age was the earliest age to assume the responsibility of helping with weekly housecleaning (such as putting away magazines, dusting furniture, and mopping floors). However, the middle-class children responded that they be-

TABLE I

MEDIAN AGES AT WHICH LOWER AND MIDDLE-CLASS CHILDREN BELIEVE THAT BOYS AND GIRLS SHOULD BE ABLE TO ASSUME GIVEN RESPONSIBILITIES

Responsibility	Lower-Class		Middle-Class	
	Boys (N = 128)	Girls (N = 124)	Boys (N = 130)	Girls (N = 126)
1. Wash their hands before each meal without being reminded	6	6	5	5
2. Take a bath without their mothers' help after they had been reminded to do so**	7.5	7	5	6
3. Keep their shoes tied*	6	7	5	6
4. Put their dirty clothes in a hamper without being told*	7	7	5.5	6
5. Hang up their coats and caps when they come from school without their mothers asking them to*	7	8	6	7
6. Put on outside coats without help from their mother*	6.5	6	5	5
7. Put their toys in their proper places without being told to do so by their mother*	6	6.5	5	5.5
8. Help with weekly housecleaning (such as putting away magazines, dusting furniture, and floors)*	8	8	7	7
9. Set the table for a family meal without help from their mother*	8	10	7	9
10. Dry the dishes (those usually used by the family)**	9	9	7	8
11. Answer the telephone and call the person that is wanted**	7	8.5	6	7
12. Take a note their mother had written to a neighbor a half-block away and come back right away**	8	8	6	7
13. Go to a familiar grocery store and choose and pay for three or four items for their mother*	8	10	7	9
14. Ride a bicycle safely in the usual amount of traffic on the streets around their home*	9	10	8.5	9
15. Ride a city bus by themselves if they were familiar with where it goes**	9	8	7	7
16. Walk to school alone if they have to cross a few moderately busy streets*	8	9	7	8
17. Cross a street that has traffic lights if they were by themselves*	8	10	7.5	9
18. Cross streets with a great deal of traffic which do not have traffic lights*	9	10	8	9
19. Sit through a church service for about an hour without bothering others*	8	9	7	8
20. Keep a younger child from walking into the street alone	9	9	8	9
21. Stay alone in their home for a half-day once in a while*	9.5	9	8	8
22. Share their wagon by taking turns with two friends without being watched by their mother**	8.5	7	6	6
23. Take something back which the borrowed from a neighborhood child without being reminded by their mother**	9	8	7	7
24. Choose guests to attend their birthday party with their mothers' help*	8	8	6.5	7
25. Go with a friend to an afternoon movie of which their mother approves*	9	10	8	9

* Significant at the .05 level. ** Significant at the .01 level.

lieved that boys and girls should be able to assume responsibility for this task by the age of seven.

Six of the differences between the responses of the lower and middle-class children related to children's relationships (Items 19, 20, 22, 23, 24, 25). Lower-class children believed that boys and girls should be able to assume the responsibility of sharing their wagon by taking turns with two friends without being watched by their mother by the age of seven or eight. The middle-class children, however, indicated that six-year-olds should be able to assume this responsibility.

Four of the differences are concerned with the care of clothing (Items 3, 4, 5, 6). The lower-class children believed that, by the age of six, boys and girls should be able to assume responsibility for putting on outside coats without their mothers asking them, whereas the middle-class children indicated that this task could be assumed by five-year-olds.

Four of the differences between the lower-class children's ratings and those of the middle class were concerned with assuming responsibility for the performance of activities alone (Items 15, 16, 17, 18). Lower-class children, for example, believed that boys and girls should be able to assume the responsibility to ride on a city bus by themselves, by the age of eight or nine, if they were familiar with where it goes. The middle-class children on the other hand, responded that seven years of age was the earliest age at which boys and girls should be expected to assume this responsibility.

Two of the differences relate to cleanliness of self (Items 1, 2). The lower-class children believed that, by the age of six, boys and girls should be able to assume the responsibility of washing their hands before each meal without being reminded, while the middle-class children indicated that five-year-old children should be able to assume responsibility for this task.

The remaining two differences between the two groups were concerned with playing alone (Items 14, 21). The lower-class children indicated that they believed that, by the age of nine, boys and girls should be able to assume the responsibility for staying alone in their homes for half a day once in a while. However, the middle-class children indicated that children eight years of age could be expected to assume responsibility for this task.

The results of this investigation thus indicate that middle-class children anticipate assuming responsibility earlier than lower-class children. This may be due to the fact that middle-class parents expect the early assumption of responsibility in the home on the part of their children. This finding corroborates the results of studies made by Ericson (2), and Davis and Havighurst (1). They found that children in middle-class families were expected to assume responsibility earlier than were the lower-class children. As Ericson (2) points out, "middle-class children are taught ways of living that will prepare them to become financially independent, to assume positions of responsibility in the home and community, and to become responsible citizens of the culture in which they live. The lower-class children are reared in families in which life is less strictly organized, and fewer demands are made upon them."

Summary

This study was designed to obtain information concerning perceptions which children hold toward the development of responsibility, and to test the hypothesis that perceptions concerning the development of responsibility are independent of social class. A Children's Responsibility Inventory, consisting of 25 items, was administered to 509 children living in Texas. A chi square analysis failed to support the hypothesis suggesting that perceptions concerning the development of responsibility are dependent upon social class. Essentially, it was found that middle-class children anticipate assuming responsibilities earlier than lower-class children.

REFERENCES

1. Davis, A. and Havighurst, R. J. "Social Class and Color Differences in Child-Rearing," *American Sociological Review*, XI (1946), pp. 698-710.
2. Ericson, M. C. "Social Status and Child-Rearing Practices," in T. M. Newcomb and E. L. Hartley, and others, *Readings in Social Psychology*, New York: Henry Holt and Company, 1947.
3. Harris, D. B. "How Student Teachers Identify Responsibility in Children," *Journal of Educational Psychology*, XLV (1954), pp. 233-239.
4. Harris, D. B. "A Scale for Measuring Attitudes of Social Responsibility in Children," *The Journal of Abnormal and Social Psychology*, LV (1957), pp. 322-326.
5. Harris, D. B. "Parental Judgment of Responsibility in Children, Children's Adjustment," *The Journal of Genetic Psychology*, LXXXXII (1958), pp. 161-166.
6. Harris, D. B. et al. "The Relationship of Children's Home Duties to an Attitude of Responsibility," *Child Development*, XXV (1954) pp. 29-33.
7. McGuire, C. and White, G. D. "The Measurement of Social Status," Research Paper in Human Development No. 3 (Revised). University of Texas, Austin, 1955.
8. Mitton, B. L. and Harris, D. B. "The Development of Responsibility in Children," *Elementary School Journal*, LIV (1954), pp. 268-277.
9. Walters, J., Stromberg, F. I. and Lonian, G. "The Development of Responsibility in Young Children," *Elementary School Journal* LIVII (1957), pp. 209-216.
10. Walters, J. *Children's Perceptions Concerning the Development of Responsibility*, unpublished paper, Florida State University, Talahassee, Florida.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 9, May - June 1963)

The Use of Intelligence Tests

RICHARD M. SANDERS
University of North Carolina

THE PURPOSE of this article is to emphasize the need for the proper use of intelligence tests. An example of the gross misuse of the Science Research Associates Primary Mental Abilities (SRA PMA) test score will be cited as an illustrative case in point.

In a recent unpublished study by the writer, one of the minor findings was a high correlation between grade point averages and PMA scores for a sample of eighth-grade students. The correlation was obtained twice on the same sample, between current grade point averages and PMA scores, and between the previous year's grade point averages and the PMA scores. The correlations were: $r = .81$ ($n = 54$) and $r = .79$ ($= 51$), respectively.

The normative data of the PMA shows a range of correlations for the various subtests and various school subjects to be from $r = .23$ to $r = .66$ with the median correlation of subtests with overall grade point averages to be $r = .43$ (Science Research Associates, 1958).

The difference between the highest correlation from the normative data ($r = .66$) and the lowest correlation found by the author's study ($r = .79$) does not quite reach significance ($p > .05$), though all other comparisons between the various subtests and various school subjects with the lowest correlation found by the author demonstrate significant differences at least at the .05 level.

The basis for the high correlations might be explained by the fact that the teachers of the school had open access to the PMA scores of their pupils, which they used frequently, seemingly for the purpose of validating their opinions of the work done by their pupils.

Four facts stand out from the above data:

1. If the hypothesis is true that knowledge of the PMA scores biased the grading of the students by the teachers, as the significantly higher than usual correlations found by the author imply, then teachers should be made more aware of the use of the PMA and like tests in order to be on guard against such biasing.

2. Directly stemming from the point made above, if there is a bias in grading produced by a knowledge of the PMA scores, then research using grades and scores as variables is obviated where this condition occurs.

3. The point seems worth making that, given the PMA scores of pupils, teachers should not be biased by these scores in their grading since the grade point average of the student is the raw data against which one might validate an intelligence test, not the data to be validated by the intelligence test.

4. The question of whether or not the teachers should have free access to the PMA scores of their students cannot be answered directly from the data presented. The answer must come from the teachers and school testing personnel who should determine whether or not a particular teacher can maintain unbiased grading of the pupils in the face of knowledge of the PMA or like scores of those pupils.

In summary, it is this writer's hope that the findings cited above and the discussion just made will motivate better control over the use of the PMA score on the part of school testing personnel, and will also make teachers more aware of the nature of tests of IQ. These tests are very important and can be very useful, not only in a school system for guidance purposes, but also in both pure and applied research; but only when used properly, not when misused as in the above case.

REFERENCE

Science Research Associates, Inc. Manual for the SRA Primary Mental Abilities, Ages 11 to 17, 1958.

Field News

THE OFFICE OF EDUCATION, HEW,

recently has issued a publication entitled Summary Report on Bachelor's and Higher Degrees Conferred During the Year 1960-61. The report shows that approximately 401,000 bachelor's and first-professional degrees, 79,000 master's degrees and 10,600 doctor's degrees were conferred during this period. The number of bachelor's degrees granted has increased during each of the last six consecutive years. However, 7.5 percent fewer bachelor's and first-professional degrees were conferred during 1960-61 than were conferred in the post World War II peak of 1949-50. The number of conferred master's and doctor's degrees is the highest ever reported.

When the 1960-61 period was compared with the 1959-60 period, foreign languages and literature, trade and industrial training, mathematical subjects and Military, Naval and Air Force Science showed the greatest increase in the percentage of bachelor's degrees conferred. During the same comparative periods, agriculture, architecture, engineering, geography, and the physical sciences had the greatest decline in percentage of bachelor's degrees conferred.

THE FORD FOUNDATION

has made recent grants totaling over $1,051,500 to four universities for the study of the problems related to the improvement of college teaching. The recipient universities are: Cornell University, Florida State University, University of Nebraska, and Washington University (St. Louis).

The programs to be carried out by the participating universities will try to strengthen the master's degree in order to make it a sound qualifying degree for college teaching. Also it is intended that the master's level work under this program would fully prepare students for doctoral work.

Some key characteristics common to the programs at all four universities are: the identification of outstanding students during the college freshman and sophomore years, acquisition of research and writing experience, the establishment of graduate courses at the junior and senior level, completion of the Ph. D. language qualifying examinations, and substantial preparation for the Ph. D. qualifying examination.

A SUBCOMMITTEE OF THE NORTH CENTRAL

Association of Colleges and Secondary Schools recently has completed a survey of secondary teachers. The primary purpose of the survey was to determine what types of human relations problems are encountered by the teachers and how well the teachers feel the college prepared them for these situations.

The committee concluded that the study yielded evidence showing that secondary teachers do not have adequate information regarding "the nature of learning and the processes of teaching, and about culture patterns of minority groups in our society," The committee also concluded that "no amount of pre-service education will be sufficient to meet the needs of the teachers." The committee felt these noted deficiencies must be corrected by the use of in-service education in the area of human relations.

THOSE PERSONS INVOLVED

in the designing of new school buildings will be interested in a new HEW publication Planning School for New Media. This guide has been prepared to assist school personnel in planning buildings so that effective use may be made of the modern instructional media.

THE NEA RESEARCH DIVISION

has computed some statistics concerning the classroom teacher based on a 1961 sample. The average teacher is 41 years of age. He has taught 13.4 years and has an annual income from teaching of $5,264. The average teacher has 4.7 years of college.

There are substantial sex differences on all these variables. The average female teacher is 8 years older, has had twice as much teaching experience, and earns $148 less per year than her male counterpart. Forty-one percent of the male teachers have a master's degree, while only 16 percent of the females have achieved this academic level.

PROJECT TALENT

has found that almost all schools use standardized tests and that an overwhelming majority plan to make greater use. As yet there is no evidence of reaction against use of standardized tests. The investigators believe that it is possible to classify high schools into 17 homogeneous categories and to develop assessment devices for valid comparisons of schools in achieving their objectives. They believe they can measure and allow for family and community situations which affect school outcomes differently. Apparently, educational outcomes are more closely related to the quality and type of students enrolled than to the school's educational policies or practices.

RETURN TO THE USE OF CORPORAL

punishment seems to be a growing trend in schools across the country as a partial answer to the problems of juvenile delinquency. This is reported in a recent research study, Effective Secondary School Discipline, by Knute G. Larson and Melvin R. Karpas (Prentice-Hall, Englewood Cliffs, New Jersey, 1963; 206 p. $4.95).

THE AVERAGE IQ OF OVER 13,000 YOUTHS

who dropped out of school in Ohio last year was estimated to be 90. In a study completed by the state department of education, two-thirds of the boys and one-third of the girls were overage for their grade level and, among the ninth grade dropouts, 97 percent were boys. E. E. Holt, Superintendent of Public Instruction, said dropouts commonly indicated that they were leaving school to enter the adult world of work and family life, but everything indicated that they were less prepared for this than their classmates who remained in school.

Readers are invited to send items of research interest to Prof. John Schmid, College of Education, University of Arkansas, Fayetteville, Arkansas, or to Dr. Willard E. North, Central Missouri State College, Warrensburg, Missouri.

Among Those Present

LEWIS R. AIKEN, JR. received his BS and MA degrees in psychology from Florida State University and his PhD in psychometrics and experimental psychology from the University of North Carolina. He was associate professor of psychology at Atlantic Christian College from 1958-60 and is currently assistant professor at the Woman's College of the University of North Carolina. Dr. Aiken has been especially interested in the effect of non-intellective factors on performance, about which he has published a number of papers. Other research and writings have been concerned with an analysis of emotion and the selection and placement of college students.

EDWARD AMIDON is assistant professor of education in the Group Dynamics Center at Temple University. He received his BS and BA degrees from the University of Minnesota and then went on to receive his MA and PhD degrees from Minnesota in educational psychology. His major research interests include applications of social psychological ideas to the classroom and to the school situation. He has worked in various workshops on human relations training and is particularly interested in the effect of human relations training on the behavior of teachers in the classroom.

JOSEPH C. BLEDSOE is professor of education at the University of Georgia. He holds the bachelor's and master's degrees from the University of Georgia and the PhD from George Peabody College for Teachers. Prior to his present position he held positions as personnel technician for the US Civil Service Commission, vocational adviser for the Veterans Administration, and educational psychologist, Air University, Maxwell Air Force Base, Alabama. During World War II, he served as personnel consultant with an Army Replacement Depot in the Pacific Area. He was a visiting lecturer in education at the University of Wisconsin during the summer term of 1958. For two years he was a Jesse Jones Fellow at Peabody College.

ARTHUR BLUMBERG is associate professor of education in the Group Dynamics Center at Temple University. He received his BS and MEd degrees from Springfield College, and his EdD from the Teachers College of Columbia University. He has had a great deal of experience in working with groups (both in education and community agencies) as a human relations and leadership training consultant. His major research interests have been concerned with group behavior in educational and community service settings. At present he is particularly interested in the relationship between certain group phenomenon and teacher attitudes toward the school, fellow teachers, and faculty meetings.

CAROL BOBBE received her BS in elementary education from New York University in 1946. She is at present completing the requirements for the master's degree in psychology at Hofstra College. Mrs. Bobbe has been employed as a research associate in Hicksville, New York since 1961.

WILLIAM J. CAMPBELL received his BA and his MA in psychology from Hofstra College. He served as an undergraduate assistant in the research bureau of the Psychology Department of Hofstra College and as a graduate assistant. He attended Purdue University on a teaching fellowship, and taught as a psychology instructor at Hofstra College, and was also a research associate at Human Resources Foundation. Now he is the director of research of the Hicksville, New York School District. Mr. Campbell has fourteen publications to date.

DONALD J. DICKINSON holds three degrees from Oklahoma State. He has been a school and hospital psychologist. In 1959-60 he was a US Public Health Fellow.

LAWRENCE M. DeRIDDER obtained his bachelor's degree at Northern Michigan College and both his master's and doctor's degrees at the University of Michigan. Between 1950-53 he was a student counselor and assistant professor in psychology at the Undergraduate Division of the University of Illinois and, from 1953 to the present, a member of the College of Education of the University of Tennessee as an associate professor of educational psychology. His articles have appeared in the Elementary School Journal, College and University, School and Society, Junior College Journal, Journal of Higher Education, and Journal of Educational Research.

CARMEN J. FINLEY is director of research in the Sonoma County Schools Office, Santa Rosa, California, a post she has held for 10 years. She obtained her master's degree from Teachers College Columbia University in 1952 and the doctor of philosophy degree in 1962 from Columbia. Prior to her present position she was a teacher of mathematics at Porterville Union High School

TRAVIS L. HAWK obtained his bachelor's and master's degree at North Texas State College and his doctorate at The University of Texas. He taught and served as elementary school principal eight years in the public schools of Texas. He has been affiliated with The University of Tennessee College of Education since 1958 as an associate professor of educational psychology.

CONSTANCE K. KAMII is a counselor at East Junior High School in Ypsilanti, Michigan, and is studying for a PhD in education and psychology at the University of Michigan. She graduated from Pomona College as a Phi Beta Kappa and received an MA degree from the University of Michigan.

J. DONALD LARSON has a BS from Northern State Teachers College, Aberdeen, South Dakota, and an MS from Northwestern. Presently he is chief of home-school counseling with the Sioux Falls, South Dakota, Schools. He has been a rural school teacher, high school principal, county probation officer, director of an institution for delinquent children, and junior high counselor.

RICHARD A. SANDERS at Dartmouth in psychology and an MA at Hollins in experimental psychology. He was employed at The MITRE Corporation, Bedford, Massachusetts, first in the Human Factors Department doing research in the general area of the physical correlates of the visual stimulus. Now he is back at Dartmouth as the assistant director of the Programmed Learning Research Project. His major interests are operant behavior, learning theory, and methods and procedures of research.

ELAINE LAMBERTI received her BS in psychology from C. W. Post College. She was awarded an academic scholarship for undergraduate work and graduated with honors. She is now attending Hofstra College where she is completing the requirements for the masters degree in psychology. She is also employed as a research associate at the Hicksville (N. Y.) School District.

JACK M. THOMPSON is director of guidance in the Sonoma County Schools Office, Santa Rosa, California. He obtained a bachelor's and master's degree from San Jose College and is presently doing graduate work at the University of California. Prior to his present position, he was a teacher of regular and special classes and a school psychologist.

CHARLES SHEPPARD graduated with honors from Hofstra College with a BA in psychology. He is presently completing the requirements for the masters degree at Hofstra. He was awarded academic scholarships for both undergraduate and graduate work. He has been employed as a research associate with the Hicksville (N. Y.) School District since 1960.

DAVID P. WEIKART is currently school psychologist and director of special services of the Ypsilanti Public Schools, Michigan. He received a BA degree from Oberlin College and an MA degree from the University of Michigan. He is at present completing his PhD dissertation on teacher attitudes towards specific educational situations.

JOHN L. HORN received his primary and secondary education in the Denver Public Schools. After serving in the army from 1950 to 1953, he completed undergraduate courses at the University of Denver. During the following year he studied under the Fulbright Plan at the University of Melbourne, returning to America and the University of Illinois in 1958 as a research assistant to Professor R. B. Cattell. He continued his graduate studies at this institution on an N. I. M. H. Predoctoral Fellowship, obtaining the MA degree in 1960 and the PhD degree in 1961. He then returned to the University of Denver as an assistant professor of psychology, the position he now holds. He has published articles, alone and in collaboration with Prof. Cattell and Dr. Scheier in the Journal of Clinical Psychology, the Journal of British Psychology, Educational and Psychological Measurement, and Genetic Psychological Monographs.

Executive Editors
WILSON B. THIEDE, Professor of Education, Educational Administration, and Agricultural and Extension Education, The University of Wisconsin, Madison, Wisconsin. Chairman.
THEODORE HARRIS, Professor of Educational Psychology and Education, The University of Wisconsin, Madison, Wisconsin.
L. JOSEPH LINS, Professor and Coordinator of Institutional Studies, The University of Wisconsin, Madison, Wisconsin.
PHILIP LAMBERT, Professor of Educational Psychology and Educational Administration, The University of Wisconsin, Madison, Wisconsin.
JOHN SCHMID, Professor of Education and Chairman of the Department of Educational Foundations, University of Arkansas, Fayetteville, Arkansas.

Editorial Board
JACOB O. BACH, Head, Educational Research Bureau, Southern Illinois University, Carbondale, Illinois.
WILLIAM A. BROWNELL, Dean, School of Education, University of California, Berkeley, California.
ROBERT A. DAVIS, Professor of Educational Research, George Peabody College for Teachers, Nashville, Tennessee.
GUY D. McGRATH, Dean, College of Education, Arizona State University, Tempe, Arizona.
HAROLD E. MITZEL, Director, Office of Research and Evaluation, Division of Teacher Education, College of the City of New York, New York City.
WILLARD C. OLSON, Dean, College of Education, University of Michigan, Ann Arbor, Michigan.
ARTHUR E. TRAXLER, Executive Director, Educational Records Bureau, New York 32, New York.

Associate Editors
HAROLD M. ANDERSON, Associate Professor of Education, University of Colorado, Boulder, Colorado.
VERNON E. ANDERSON, Dean, College of Education, University of Maryland, College Park, Maryland.
RUDYARD, K. BENT, Professor of Education, University of Arkansas, Fayetteville, Arkansas.
EMMETT ALBERT BETTS, Research Professor, University of Miami, Coral Gables, Florida.
LAWRENCE P. BLUM, Assistant Director, Student Personnel Services, University of Wisconsin–Milwaukee, Milwaukee 3, Wisconsin.
J. MURRAY LEE, Chairman, Department of Elementary Educa-
ion, Southern Illinois University, Carbondale, Ill.
. WELTY LEFEVER, Professor of Education, University of
outhern Caifornia, Los Angeles, California.
. E. TORSTEN LUND, Associate Professor of Education, Uni-
ersity of California, Berkeley, California.
VILLIAM A. McCALL, Lecturer in Education, University of
Miami, Coral Gables, Florida.
ELYTE W. MORRIS, President, Southern Illinois University,
arbondale, Illinois.
ALWORTH PLUMB, Chairman, Division of Education and
sychology, University of Minnesota (Duluth Branch), Duluth,
innesota.
AVID G. RYANS, Rand Corporation, Santa Monica, California.
ESTER B. SANDS, Professor of Education, University of
alifornia, Santa Barbara, Goleta, California.
ANDEL SHERMAN, President, and Consultant, Emotional
ealth Association, Los Angeles, California.
ILDRED B. STANTON, Consultant, Special Education, State
epartment of Education, Hartford 15, Conn.
. W. THEISEN, Professor of Education, Marquette University,
iwaukee, Wisconsin.
. R. TRABUE, Professor of Higher Education, University of
entucky, Lexington, Kentucky.
WAYNE WRIGHTSTONE, Director, Bureau of Educational
esearch, Board of Education of City of New York, Brooklyn,
ew York.

ditors Emeriti
J. ASHBAUGH, Dean Emeritus, School of Education, Miami
iversity, Oxford MR50, Ohio.
RRY J. BAKER, 19050 Wiltshire, Lathrup Village, Michigan.
WLER D. BROOKS, Director Emeritus, Department of Psy-
ology, DePauw University, Greencastle, Ind.
O J. BRUECKNER, Professor Emeritus, University of Minne-
a, Minneapolis, Minnesota.
R. BUCKINGHAM, 33 Chatham Circle, Wellesley Hills,
ssachusetts.
AR A. DOLL, Director Emeritus, Devereux Schools, Devon,
nsylvania.
RL R. DOUGLASS, Director Emeritus, School of Education,
versity of Colorado, Boulder, Colorado.
DUNCAN GRIZZELL, Dean Emeritus, University of Penn-
ania, 640 Maxwelton Court, Lexington, Ky.
NEST HORN, Professor Emeritus of Education, State Uni-
ity of Iowa, W307 East Hall, Iowa City, Iowa.
RL HUDELSON, Professor Emeritus of Education, West Vir-
a University, Morgantown, West Virginia.
M. JORDAN, Professor Emeritus, University of North Caro-
Chapel Hill, North Carolina.
RED S. LEWERENZ, Director Emeritus of Evaluation and
arch, 4107 Holly Knoll Drive, Los Angeles California.
R. MEAD, Director Emeritus of Educational Research, Uni-
ity of Florida, 1719 N.
J. NELSON,
3 West Or

The Journal of Educational Research

Dedicated to the Scientific Study of Education

VOLUME 56 JULY-AUGUST 1963 NUMBER 10

It Says in the JER ---------- 506

Screening in Schools for Behavior Disorders: Use of Mothers' Reports of Symptoms—*John C. Glidewell, Herbert R. Domke, Mildred B. Kantor* ---------- 508

A Comparison of the Common Denominator and Inversion Methods in Teaching Division of Fractions—*Lelon Roger Capps* ---------- 516

Elementary School Teachers Rate Report Cards—*Norman M. Chansky* ---------- 523

The Outcomes of Curriculum Modifications Designed to Foster Critical Thinking—*Norman E. Wallen, Vernon F. Haubrich, Ian E. Reid* ---------- 529

Typing Performance as Related to Mental Abilities and Interests: A Preliminary Study—*Hortense G. White* ---------- 535

Comparison of Student Reaction to Similar, Concurrent Teacher-Content Instruction—*Shirley Ullman Wedeen* ---------- 540

An Analysis of Remedial Activities Used by Elementary Teachers in Coping with Classroom Behavior Problems—*Donald L. Barnes* ---------- 544

A Comparison of the Performance of Pre-Graded Students with Grade-Motivated Students—*T. L. Hawk* and *L. M. DeRidder* ---------- 548

The "Successful" Child: Some Implications of Teacher Stereotyping—*Philip Lambert* ---------- 551

Six-Grade Children Evaluate Twelve Cases of Malicious Mischief—*Alvin E. Robinson* ---------- 554

Six-Grade Ch chief—*Alvin E. Robinson* ---------- 556

Field News ---------- 556

Among Those Present ---------- 559

Index

The Journal of Educational Research is published at Madison, Wisconsin, monthly 10 times a year (the May-June and July-August issues being combined). The subscription price is $7.50 a year, plus $1 postage for all subscriptions outside the continental United States. Single copies, $1. Second class postage paid at Madison, Wisconsin, Copyright 1963 by Dembar Educational Research Services, Inc.
ndence care of DERS, Box 1605, Madison 1, Wisconsin.
ARCH SERVICES, Inc. WALTER FRAUTSCHI, President.
ENCE A. SCHOENFELD, Assistant

It Says in The JER...

CAN YOU RELY ON MOTHERS' REPORTS IN SCREENING SCHOOL-AGE CHILDREN FOR BEHAVIOR DISORDER ? Yes, says a report from the St. Louis (Missouri) County Health Department. A reliable, positive relationship seem to exist between the number of symptoms reported by a child's mother and the child's degree of maladjustment as reported by the child's teacher, so a symptom inventory based on mothers' reports is a useful and economical screening device.

WHAT GOOD ARE GRADES? They are a spur to class performance, according to a study at the University of Tennessee. A group of students were awarded grades in a course at the beginning of the term on the basis of previous grade-point averages. Compared to a control group, removal of pressure for a grade acted as a depressant on work output and on test scores.

ALMOST EVERY INSTRUCTOR some time in his teaching career has had the feeling that while he has attempted to give the same instruction to two or more sections of the same course, as the semester moves along he seems to be giving two different courses. Does he imagine the feeling or can different groups of students respond differently to the same presentation? A Brooklyn College study suggests that identical assignments and exams can be perceived quite differently by different groups of students.

IS GOOD BEHAVIOR IN CLASS more important than teachers think in determining whether a child will be considered successful? Does the elementary school teacher somehow tend to discriminate against boys? What sorts of reinforcement in the form of teacher approval are given to children considered by the teacher to be successful? These and other penetrating questions are raised by Lambert in a pilot study of the relationship of teacher stereotyping to pupil performance.

STUDENTS WHO DO WELL DO WELL IN TYPING COURSES may be characterized as (1) doing well in first-year general high school average, (2) doing well in math courses during the first year, (3) performing well on general intelligence tests, and (4) having scientific interests. So reports White from Nashville.

THREE UNIVERSITY OF UTAH RESEARCHERS report favorably on the outcomes of a curriculum modification designed to foster critical thinking through acquainting students with the principles of logic and experimentation.

IN TEACHING DIVISION OF FRACTIONS, which method is best, the common denominator method or the inversion method? Capps reports that neither method is superior to the other in terms of teaching division, but that the common denominator method seems to reinforce skill in multiplication of fractions.

HOW SOUND ARE REPORT CARDS? As a concise summary of a child's progress in school, some types of report cards fall short of the mark, while others convey to the parent functional information, Chansky reports from a New York student.

WHEN ASKED TO IDENTIFY FACTORS UNDERLYING CLASSROOM BEHAVIOR PROBLEMS, many Colorado teachers were unable to do so, and they frequently took no remedial action of any kind, Barnes reports. He suggests much more needs to be done to help elementary teachers study and interpret behavior problems of elementary-age children.

WHEN ASKED TO REACT TO COURT CASES INVOLVING MALICIOUS MISCHIEF, a group of California sixth-graders displayed little congruence with the court's interpretation of the right of acting. Robinson suggests the school has a responsibility to help children develop value judgments appropriate to a democratic society.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 10, July-August 1963)

Screening in Schools for Behavior Disorders: Use of Mothers' Reports of Symptoms

JOHN C. GLIDEWELL
HERBERT R. DOMKE
MILDRED B. KANTOR
St. Louis County Health Department
Clayton, Missouri

SCREENING HAS long been a basic component of school health services. School mental health services, however, have not yet found a simple and reliable method for screening school-age children for behavior disorders.

The evaluation of preventive mental health programs requires a measure of adjustment which can be used with economy with large groups of subjects. As a part of a larger study designed to evaluate a community mental health program (2), several sub-studies have been undertaken to test the screening efficiency of a symptom inventory designed for use with mothers. A small-sample study has been reported previously (3). This is a report of the first large-sample study of the screening method.

Whether knowledge about symptoms alone can be used as a screening tool depends on whether they can be used successfully to predict the presence and degree of clinical sickness in children. This subject has been discussed among child psychiatrists for many years. Some feel it is dangerous to use presence or absence of symptoms as an indication of whether the child is sick or well. Gardner (4) is a representative of this group. He divides children's problems into two categories, i. e., those requiring short-term therapy, and those requiring long-term treatment. Of the short-term group he says that therapy aimed at relieving the symptom "will probably allow the child an extremely good chance of continuing thereafter a normal and orderly personality development," (4, pp. 131-33). Of the long-term group he warns against using absence or disappearance of symptoms as an indication either for stopping treatment or of the recovery of the child. He says, (4, p. 133), "Removal of the presenting symptom may delay but not prevent the child from developing a more serious disability later in childhood or in adult life." In summary, Gardner, appears to doubt that symptoms themselves can be used to discriminate between sick or well children, or those with different degrees of emotional adjustment.

On the other hand, Levy believes that symptoms can be used to determine sickness or health in children. He says, (11, p. 227), "One of the findings that resulted from a series of follow-up studies starting at the Institute of Child Guidance in New York and utilizing material from about a dozen child guidance clinics . . . was that the highest correlation with general personality growth was the disappearance of the presenting symptoms." Witmer (16) also finds that disappearance of symptoms in children treated in a child guidance clinic correlates higher with clinical improvement than any other factor.

Support for the use of symptoms in determining adjustment in adults is found in recent work by Hunt, Wittson, and Hunt (8, 9) where there was demonstrated to be a positive relationship between the number of presenting symptoms in adults, the severity of the neuropsychiatric disorder, and suitability for military service.

Kanner (10) lays much emphasis on the "complaint," i. e., the mother's statement of symptoms, as a critical factor in determining treatment. He says, "The clearing up, or persistence, of the difficulties complained of may frequently, though not always, be used as a measure of the reasonableness and adequacy of the handling of the maladjustment."

In 1954, Eitzman (7) found, in an intensive follow-up study of twelve children treated in the St. Louis County Child Guidance Clinic, that mothers reported markedly fewer symptoms after treatment than they reported at the time the child was referred. Specifically, at the time of referral, the twelve mothers reported a total of 86 current symptoms (mean, 7. 2). After treatment, these same mothers reported a total of only 40 symptoms (mean, 3. 3).

In a study of 91 public school children and 35 children referred to a child guidance clinic, the St. Louis County research group, in 1955 (7), found that a reliable, positive relationship exists between the number, frequency, duration, and severity of the symptoms reported by a child's mother and the degree of sickness found in the child.

The Methods of the Present Study

Purpose. The present study is a large-sample extension of previous work in St. Louis County (7). Its major purpose is to test the effectiveness with which a mother's symptom report can be used to identify, in advance, those children who would be referred by a teacher as at least sub-clinically disturbed.

The sample. The sample selected for study (6) is composed of 830 white families, each having at least one child in the third grade in one of fifteen schools located in St. Louis County. The areas served by the schools vary widely with respect to social characteristics. The schools were selected on the basis of willingness of the school administration to participate in the research program and on their ability to supply, from a single school, two third-grade classrooms necessary for the research design.

The third grade was selected because at this grade level the children were expected to have relatively few transient symptoms of disturbance. In the third grade, most children have completed the initial behavioral adjustments to the move from home to school. At the same time, they have not yet begun to make the transition to pre-adolescence and adolescence.

Criteria of adjustment. In the St. Louis County school mental health services, as in many other such services, a child is diagnosed as disturbed when he has been 1) referred by the teacher, and 2) seen and accepted as disturbed by professional mental health personnel (1, 6). Under these conditions, the first step in effective screening is the identification of the children who will be referred by the teacher.

For this study, ratings of general adjustment were made by the children's teachers on a four-point rating scale. The adjustment categories were derived from those employed in an early study by Ullmann (14).

1. Well adjusted: A happy child who is well adjusted in his relationships with others and in his accomplishments.
2. No significant problems: A child who gets along reasonably well and accomplishes reasonably well the things that usually go with his age and level of development.
3. Sub-clinically disturbed: A child who is not so happy as he might be; has moderate difficulties getting on; growing up represents something of a struggle.
4. Clinically disturbed: A child who has, or, at his present rate is likely to have, serious problems of adjustment, and needs clinical help because of such problems.

Six percent of the sample (n=49) were given ratings by psychiatric social workers based on professional findings. These independent ratings were exactly alike for 53% of the 49 cases and within one point in 92% of the cases. There was 80% agreement as to whether or not a child showed at least a mild disturbance. This same trend appeared in the previous study (2), where it was found that the teachers' opinions about the emotional state of the children agreed exactly with the opinion of the professional mental health personnel in 86% of the cases.

Such professional ratings would permit a second, and a more effective step, in screening--the identification of all children who would be accepted by clinicians as disturbed, whether or not they would be referred by the teacher. The work reported here, however, was directed toward the first, more limited, step of predicting in advance which children would be referred by the teacher. In later work, the screening instrument can be checked against the diagnosis of professional mental health personnel (5). This work must await the time when a significant sample of children has been seen by professionals.

For the present purposes, the children were classified into one of the four adjustment levels according to the professional rating where the professional personnel knew the child and supplied a rating for him. The teacher's rating was used where a professional rating was not available. By this method of assessment, the sample of 830 included 167 well adjusted children, 429 children with no significant problems, 166 sub-clinically disturbed children, and 68 clinically disturbed children--or about 28% with at least sub-clinical disturbance.

The interview. The information about the presenting symptoms of the child was collected through a private, home interview with the mother of the child. The mother was asked a series of questions about various characteristics of the family and the family background. Following this she was introduced to a series of questions which constitute the symptom inventory and take this form: "Does Johnny have any trouble.... (eating, sleeping, getting along with other children, etc)?" Affirmative responses were followed by probes into a) the specific nature of the difficulty, b) the frequency, c) the duration, and d) the severity. This line of questioning was continued through twenty-one areas of difficulty:

a. eating
b. sleeping
c. digestion
d. getting along with children-acting out

e. getting along with children-withdrawn
f. getting along with adults-acting out
g. getting along with adults-withdrawn
h. unusual fears
i. nervousness
j. thumbsucking
k. overactivity
l. sex
m. daydreaming
n. temper tantrums
o. crying
p. lying
q. stealing
r. destructiveness
s. rejection of school
t. wetting
u. speech

The number of symptoms is determined simply by counting the number of symptoms reported by the mother. The range of possibility extends from the minimum of reporting no presenting symptoms in the child to the maximum of twenty-one. In this study, 129 mothers reported no symptoms, the mode was two, and only one mother reported as many as twelve symptoms.

Findings

Substantiation of the screening effectiveness of the symptom inventory depends on two sorts of findings: 1) that a statistically significant relationship exists between the teachers' ratings of adjustment and the number of symptoms reported by the mother; 2) that the nature of the relationship will permit an accurate prediction of disturbance based on the screening findings alone.

The existence of the relationship. The data in Figure 1 and Table I illustrate the linear relationship found between the number of symptoms reported and the ratings of adjustment. Mothers of children without disturbance reported, on the average, two symptoms. Mothers of disturbed children reported three or more symptoms.

An analysis of the variance in the number of symptoms is shown in Table II, and it demonstrates that the relationship found here could have occurred by chance less than one time in a thousand. The analysis shows also that, in spite of differences in adjustment ratings in the two sexes, there are no significant differences in the number of symptoms reported for boys and girls. The general relationship between number of symptoms and disturbance holds specifically for boys and girls, taken separately.

The same analysis of variance has been accomplished separately for four socio-economic groups--a) upper, b) upper-middle, c) lower-middle, and d) lower (15). The existence of the general relationship is confirmed specifically for each of the four groups.

These findings confirm the existence of this relationship as found on smaller samples in earlier work (7). The current large-sample findings, however, indicate a more restricted range of differences among the means of adjustment groups (0.0 to 4.8) than did the previous findings (1.5 to 7.4).

The findings presented thus far show that there is a reliable and positive relationship between the number of symptoms reported by a child's mother and the degree of maladjustment reported by a child's teacher. The fact that the relationship exists and that the correlation is not perfect means that one must specify the relative accuracy with which one may predict from the symptom inventory to the teacher's rating. The accuracy of such prediction is the measure of effectiveness of the symptom inventory as a screening device in school mental health. The discussion to follow will be devoted to an evaluation of the symptom inventory as a screening instrument, under the conditions imposed by two of the most common approaches to public health screening --1) case finding, and 2) screening for referral to limited facilities.

Screening for case finding. To help explain one approach to evaluating the symptom inventory, one might review another more familiar screening device. The 70 mm x-ray is used in tuberculosis screening primarily to identify and select a group of people who may possibly be diseased and who may require further study. For this screening to be effective, the group selected for further study must include almost all the actually diseased people. It may, and usually does, include also a very high proportion of disease-free people.[2] Thus, the 70 mm x-ray yields a very high false-positive rate --the group selected for further study usually contains about 90% disease-free people. This high false-positive rate is of little concern, because the false-positive decisions can be so readily corrected by 14 × 17 x-ray results. The primary purpose of 70 mm screening is to reduce what would otherwise be a large and expensive load on 14 × 17 x-ray facilities.

This represents one approach to screening--an approach which is primarily aimed at selecting a group of people who require further study. The most important task of such screening is to select a group which contains almost all the actually diseased people. The most important criterion for evaluating such screening is the extent to which actually diseased people are missed; that is, not included in the group selected for further study. These are the cases erroneously labeled "negative" by screening; they are, in fact, diseased. These are the false-negatives, or missed cases, and the missed-case rate is the measure of effectiveness for such an approach to screening as this.

We may take this case-finding approach in mental health screening. We may decide that we will be satisfied with a screening technique which identifies those children who require further study and which "misses" a minimum of disturbed children --has a minimum missed-case rate. For this approach we need to decide just what percent of valid cases we are willing to miss. To make such a decision, we must decide how many symptoms constitute sufficient evidence of the need for further study. In short, we must select a cutting point.

TABLE I

DEGREE OF ADJUSTMENT AND MEAN NUMBER OF SYMPTOMS BY SEX

	Mean Number of Symptoms by Sex		
Degree of Adjustment	Boys	Girls	Total
Well Adjusted Children	1.7	1.7	1.7
Children with no Known Problems	2.0	2.0	2.0
Sub-clinically Disturbed Children	2.6	2.7	2.7
Clinically Disturbed Children	3.0	3.1	3.0
Grand Means	2.1	2.1	2.1

TABLE II

ANALYSIS OF VARIANCE FOR TOTAL NUMBER OF SYMPTOMS REPORTED
(Log Transformation; Corrected for Disproportionality)

Source of Variance	Degrees of Freedom	Sum of Squares	Mean Squares	F	P
Degree of Adjustment	3	1.75	0.58	11.77	.001
Sex	1	0.00	0.00	--	ns*
Interaction	3	0.00	0.00	--	ns
Residual	822	40.77	0.05		

* not significant

TABLE III

CUTTING POINTS, MISSED CASE RATES, AND SELECTION RATES

Cutting Point Between (Number of Symptoms)	Missed Case Rate	Selection Rate
0-1	09	84
1-2	26	63
2-3	47	42
3-4	60	26
4-5	74	16
5-6	82	10
6-7	88	06
7-8	91	03
8-9	95	02
9-10	97	01

TABLE IV

CRITERIA FOR PREDICTING EFFECTIVENESS OF THE SYMPTOM INVENTORY

Cutting Point Between (Number of Symptoms)	Missed Case Rate	False Positive Rate	Percent Valid Decisions	Selection Rate
0-1	09%	82%	38%	84%
1-2	26%	60%	50%	63%
2-3	47%	37%	60%	42%
3-4	60%	21%	68%	26%
4-5	74%	12%	70%	16%
5-6	82%	08%	71%	10%
6-7	88%	04%	73%	06%
7-8	91%	01%	73%	03%
8-9	95%	00%	73%	02%
9-10	97%	00%	73%	01%

Cutting points and predictive efficiency. The symptom inventory gives a simple count of the number of symptoms reported. For this sample, the count ranges from zero to twelve. From this range of scores, we need a cutting point--a score above which a prediction of possible disturbance may be made. We could select any point between zero and twelve, but the best cutting point is the one yielding the lowest missed-case rate.

The data are presented in Table III, and they illustrate a property of imperfect screening tests: the missed-case rate varies with the cutting point. In this case, the missed-case rate increases as the cutting point increases from zero to ten. The lowest missed-case rate is at zero. According to the reasoning thus far, the best cutting point is between zero and one, because it is at this point that the fewest cases are missed. This conclusion is deceptively simple, however, and further examination of the data will produce a significant qualification to the conclusion.

The data in Table III also show the proportion of the total sample selected by screening for further study. For this instrument, it is easy to see that, if only one symptom is taken as evidence of the need for further study (cutting point between zero and one), very many children will be selected for such study. Similarly, very many valid cases will be selected and only a few cases will be missed. In this case, 84% of the population is selected for further study, and, by chance, we would expect 84% of actually diseased children to be so selected. The fact is that 91% of actually diseased children are selected.

Further, when so many children are selected for further study, the screening instrument makes but little reduction of the load placed on the facilities used for carrying out the further study. Such a consideration brings out the modification which must be added to the previous conclusion about selecting the cutting point with the lowest missed-case rate. The best cutting point must satisfy not just one, but at least two criteria: 1) a minimum missed-case rate, and 2) a minimum selection rate --the percent selected for further study should exceed the prevalence rate by a minimum amount. As shown in Table III, the missed-case rate and the selection rate are inversely related. If we reduce the missed-case rate, we must increase the selection rate.

Applying a common criterion to the evaluation of missed-case rates, we may say that the highest acceptable missed-case rate would be that obtained by an unguided, chance determined, random selection for further study. These and other data lead to an estimate that the white public school children in St. Louis County are similar to those elsewhere (14), in that about 30% show at least mild disturbance. Suppose that, by random selection, a sample was divided so that 30% were called positive and 70% called negative. Because 70% of all children would be called negative, the same proportion--70%--of actually disturbed children would be called not disturbed, and that would be the missed-case rate expected by chance--70%.[3]

Returning to the data in Table III, we see that any cutting point below four would give an acceptable missed-case rate as compared to chance findings.

To evaluate selection rates, we may say that, because 30% of the public school children show disturbance, any selection rate less than 30% would most certainly miss cases. The data in Table III show that we must select a cutting point below three in order to yield an acceptable selection rate, and the problem is narrowed to selecting one of the first three cutting points.

We may now apply more practical, non-statistical criteria. A cutting point above two symptoms would miss 47% of the actual cases. Such results are statistically an improvement over chance, but any screening which misses almost half the cases is, practically speaking, pretty poor screening. A cutting point above one symptom would usually be considered a better one. It would miss only 26 percent of the cases, and it would select 63% of the total children screened and refer them for further study.

All of the foregoing analyses may be summarized by saying that the symptom inventory can be demonstrated to be 74% to 91% effective in case-finding, depending upon the availability of facilities for further study.

Screened for limited facilities. "Further study" can mean very limited or very extensive psychiatric study. It may be that facilities are not available for even limited further study of any major portion of the children screened. Under such conditions, it becomes important that 1) fewer children be selected for study, and 2) a very high proportion of those children selected be actually disturbed.

Mental health screening may differ from the more familiar tuberculosis screening with respect to the facility for "further study." The group selected by 70 mm x-ray could contain 90% disease-free people because these false-positives are so easily corrected by 14 × 17 x-ray results. Further psychiatric study is not so readily available, and we may be required to reduce the number of false-positives. Even limited psychiatric study is expensive and cannot be so justifiably expended on disease-free people as can 14 × 17 x-ray.

Under these conditions, our symptom inventory must satisfy three criteria: 1) minimum missed-case rate, 2) mimimum false-positive rate, and 3) optimal selection rate (i.e., adjusted to the facilities available). The first two criteria can be simplified. A minimum of missed cases means a maximum of valid decisions that a child is not disturbed. A minimum of false-positives means a maximum of valid decisions that a child is disturbed.

The two simply mean that we make a maximum of valid decisions, whether the decision be that the child is disturbed or that he is not.

It should be made explicit that the proportion of total possible valid decisions is, in part, determined by the prevalence rate of the disease to be identified. For example, if a disease occurs in only 1% of a population, a random selection of 1% as diseased will yield 98% correct decisions. This success is due to the simple fact that the disease occurs in only 1% of the population. In our case, the prevalence rate is 30%, and random decisions will be correct 58% of the time. Thus, we may say that the percentage of correct decisions made on the basis of the symptom inventory must be significantly greater than 58% in order to give any improvement over the results possible without any screening instrument.

Table IV presents the data, and it shows, for each cutting point, the missed-case rate, the false-positive rate, the percentage of valid decisions, and the selection rate.

Table IV shows that in this case (although this is not true of all screening), the percentage of valid decisions rises with the cutting point and reaches a maximum of 73% at six symptoms. It exceeds the minimum acceptable point of 58% at two symptoms. At two symptoms 68% are selected--a high rate for limited facilities. At five symptoms the selection rate is 10%, probably small enough for a limited facility. Only 8% of the disease-free are selected as positive, and about half of the group selected for further study prove to be, in fact, diseased. All of these factors are suitable to limited facilities, but the price is paid in missed-cases--82% of actually disturbed children are missed--that is, not selected for further study.

This analysis may be summarized by the statement that the symptom inventory can be used with 71% success to select a small number of children for further study by a limited facility.

Discussion of the Findings

The foregoing presentation of the results of statistical tests leaves several questions open for discussion: 1) How confident can one be that there is, in fact, a relationship between mothers' reports of symptoms and teachers' ratings of adjustment? 2) Should one begin immediately to use the symptom inventory as a screening instrument?

The first question about the possible relationship between mothers' reports and teachers' ratings is given as clear-cut an answer as correlational research can give. The existence of the relationship was demonstrated in a pilot study; it was confirmed in this large-sample study; and it has been confirmed on several sub-divisions of the large sample. The several sub-divisions confirm its existence separately for both boys and girls, and separately for four levels of socio-economic groups. It appears quite clear that, at least for samples in St. Louis County, the greater the number of symptoms the mother reports, the greater is the likelihood that the child will be rated as disturbed by the teacher. If confirmation is obtained for the indicated agreement between teachers' ratings and clinical examination, the data will then confirm a relationship between mothers' reports and the existence of clinical sickness in children.

With respect to the second question about screening effectiveness, the current findings are quite comparable to previous small-sample findings. The small sample showed a 93% success (7% missed-case rate) in case-finding (7, p. 53), with a cutting point after one symptom. The selection rate at this point was 75%. For comparison, the selection rate should be kept constant, and the best available comparison is with the 84% selection rate obtained by cutting after zero in this large-sample study. The findings are quite similar--93% in the small sample; 91% in this large sample.

The total valid decisions reported for the small sample in screening for limited facilities was 67%, cutting after three symptoms (7, p. 52). The current study gives 68% valid decisions at the same cutting point. Altogether, these findings indicate screening efficiency better than many psychiatric screening techniques, the general effectiveness of which has been reviewed by Meehl (12, 13).

Precise comparison to 70 mm x-ray screening in public health is complicated by the markedly lower prevalence of tuberculosis. With a prevalence of one in a thousand, one could select for further study 100 times as many as are actually diseased, and one would still be selecting only 10% of the group screened. This leaves a lot of room for error--you would have to "hit" only one time in a hundred. With a 30% prevalence rate in mental health, the selection by screening of only three times as many as are actually sick gives a 90% selection for referral to further study--you would have to "hit" one out of three. Thus the low prevalence of tuberculosis allows for a smaller selection and more error in those selected, and this is, in fact, what one finds. One in ten of those selected by 70 mm x-ray are valid positives; one in two of those selected by the symptom inventory are valid positives.

On the other hand, the 70 mm x-ray misses only one-half of one percent of cases--actually diseased individuals. The symptom inventory misses, at best, nine percent of the cases. Thus it is that mental health cases are harder to identify, even though they occur with greater frequency. This is altogether in accord with the quite limiting fact that diagnosis in mental health is still relatively quite inexact, even with the most elaborate of clinical resources.

Taking into account the efficiency of other psychiatric screening techniques, the prevalence rate of mild disturbance in public schools, and the state of the science in psychiatric diagnosis, the symptom inventory tested in this study compares quite favorably with other screening techniques.

Summary

Data have been presented to show that, in a sample of 830 white, third-grade, public school children, a reliable, positive relationship exists between the number of symptoms reported by a child's mother and the child's degree of maladjustment as reported by the child's teacher. As a case-finding instrument, a symptom inventory based on mothers' reports can be used with 74% effectiveness to screen and select for further study, children who will be rated as disturbed by their teachers--provided that facilities are available for further study of about 60% of the children screened. The symptom inventory can be used with 71% effectiveness to select a small number of children for further study by a limited facility.

Taken altogether, the findings indicate that a mother's report of symptoms constitutes a useful gross index of a child's mental health from the point of view of teachers and mothers. As such, it is a useful and economical tool for evaluating mental health for large groups of children. As a screening device, the symptom inventory compares quite favorably with other screening devices.

FOOTNOTES

1. Paper presented before the joint meeting of the Maternal and Child Health, Mental Health, and School Health Sections of the American Public Health Association and the American School Health Association at the 85th Annual Meeting of the Amer. Public Health Ass'n., Cleveland, Ohio, November 12, 1957.

   This is a report on a project of the Mental Health Research Program of the St. Louis County Health Department, supported by research grant M-592 from the National Institutes of Health, U. S. Public Health Service.

2. To establish clarity of terminology the following definitions are made explicit:

   a. Selected for further study is a phrase used to identify that group of individuals whose screening test results are positive and who may or may not be actually diseased and who may or may not be actually diseased, as would be determined by competent diagnosis. They are sometimes labeled simply "selected," "screened positive," or "identified for further study."

   b. Actually diseased is a term used to identify individuals who are actually diseased, as determined by competent diagnosis.

   c. Disease-free is a term used to identify those individuals who are free of the disease in question, as determined by competent diagnosis.

   d. False positives is a term used to identify those individuals who are selected for further study—screened positive—but who are, in fact, disease-free.

   e. False negatives or missed cases are terms used to identify those individuals who are not selected for further study—screened negative—but who are, in fact, diseased.

3. These statements of expected chance findings assume a selection rate of 30%, based on the prevalence of the disease to be identified. The expected missed-case rate would vary with the selection rate. It would, therefore, be different for other selection rates and, in the same way, it is different for each cutting point in Table III. All of the missed-case rates in Table III are at least five sigma units lower than the rate expected by chance, assuming the selection rates to be those in the last column of the Table. Many of the selection rates, however, represent inneffective screening.

REFERENCES

1. Buchmueller, A. D., and Gildea, Margaret, C. L. "Group Therapy for Parents of Behavior Problem Children in Public Schools: Paper no. 4'', International Journal of Social Psychiatry, (1956) 1(3).
2. Domke, H. R. "An Evaluation of a Preventive Community Mental Health Program: Progress Report III." St. Louis County Health Department, Clayton, Missouri, (Dec. 1956).
3. Eitzman, Conita C. "A Follow-up Study of Children Treated at the St. Louis County Guidance Clinic Between Oct. 1, 1950 and Oct. 1, 1952." Unpublished MSW Thesis, St. Louis University, (1954).
4. Gardner, G. Psychiatric Treatment (New York: Williams and Williams, 1953).
5. Gildea, Margaret, C. L., Domke, H. R., Mensh, I. N., Buchmueller, A. D., Glidewell, J. C., Kantor, Mildred B. "Community Mental Health Research: Findings After Three Years." American Journal of Psychiatry, in press.
6. Glidewell, J. C., Mensh, I. N., Domke, H. R., Gildea, Margaret C. L., and Buchmueller, A. D. "Methods for Community Mental Health Research." Americal Journal of Orthopsychiatry, (1957) 27(1)pp. 38-54.
7. Glidewell, J. C., Mensh, I. N., and Gildea, Margaret C. L. "Behavior Symptoms in Children and Degree of Sickness," American Journal of Psychiatry, (1957) 114(1) pp. 47-53.
8. Hunt, W. A., Wittson, C. L., and Hunt, Edna B. "The Relationship Between Amount of Presenting Symptomatology and Severity of Disability," Journal of Clinical Psychology,

(1955)11(3), pp. 305-6.

9.Hunt, W. A., Wittson, C. L., and Hunt, Edna B. "The Relationship Between Definiteness of Psychiatric Diagnosis and Severity of Disability." Journal of Clinical Psychology, (1952)8(3), pp. 314-15.

10.Kanner, L. Child Psychiatry (Baltimore, Charles C. Thomas, 1942).

11.Levy, D. M. Maternal Overprotection. (New York: Columbia University Press, 1947)

12.Meehl, P. E. Clinical Versus Statistical Prediction. (Minneapolis: University of Minnesota Press, 1954).

13.Meehl, P. E., and Rosen, A. "Antecedent Probability and the Efficiency of Psychometric Signs, Patterns, or Cutting Scores, Psychological Bulletin, (1955), LII, pp. 194-216.

14.Ullmann, C. A. "Identification of Maladjusted School Children." Public Health Monograph, no. 7 (1952).

15.Warner, W. L. Social Class in America. Chicago: Science Research Association (1949).

16.Witmer, Helen. Smith College Studies in Social Work, III (1933).

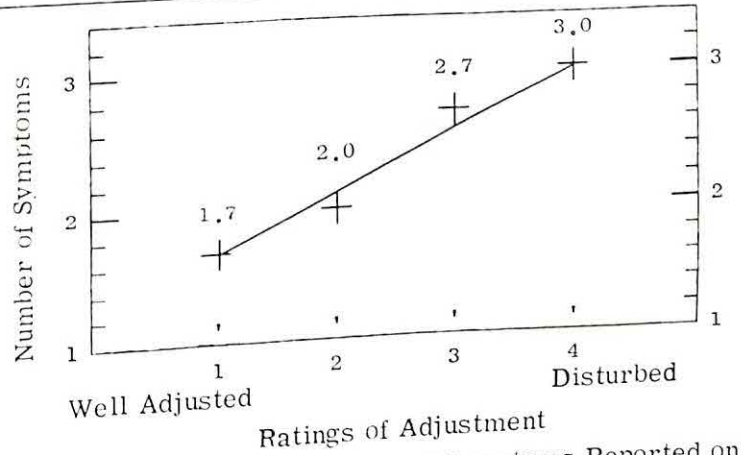


Figure 1. Regression of Mean Number of Symptoms Reported on Degree of Adjustment in the Child.

A Summary of Findings

In a sample of 830 white, third-grade, public school children, a reliable, positive relationship exists between the number of symptoms reported by a child's mother and the child's degree of maladjustment as reported by the child's teacher. As a case-finding instrument, a symptom inventory based on mothers' reports can be used with 74% effectiveness to screen and select for further study, children who will be rated as disturbed by their teachers-- provided that facilities are available for further study of about 60% of the children screened. The symptom inventory can be used with 71% effectiveness to select a small number of children for further study by a limited facility.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 10, July-August 1963)

A Comparison of the Common Denominator and Inversion Method in Teaching Division of Fractions[1]

LELON ROGER CAPPS
University of Kansas

THERE EXISTS at present a difference of opinion with regard to the effectiveness and values of the inversion and common denominator methods of teaching division of fractions. In general proponents of either of the two methods have based their decisions on limited information primarily because of the lack of a substantial amount of research to confirm either view.

Probably because it lends itself readily to a mechanical type of teaching, the majority of elementary-school textbooks have favored the inversion approach in the past. Current trends in "meaningful" teaching have lead some authorities to hypothesize that the common denominator method may secure better results because, from their viewpoint, it is more easily rationalized. The study reported here was designed to provide added information on the question concerning the effectiveness of the two methods in teaching the division of fractions.

This investigation attempted to determine if the use of either of the two methods would demonstrate significantly better results in the areas of initial achievement and retention.

The general hypotheses to be tested were:

1. There are no significant differences between methods in the area of initial achievement in division of fractions as measured by a computational test.
2. There are no significant differences between the two methods in the effect they demonstrate on skills in addition and subtraction of fractions and multiplication of fractions.

Definitions of terms used:

Inversion Method (Control)

The inversion method of division of fractions is performed by inverting the divisor and multiplying the numerators and denominators.

Example: $3/4 \div 1/2 =$
$3/4 \times 2/1 = 6/4 = 1\frac{1}{2}$

Common Denominator Method (Experimental)

The common denominator method of division of fractions is performed by changing the fractions involved to the same terms and then taking the quotient of the numerators and denominators.

Example: $3/4 \div 1/2 =$

$$3/4 \div 2/4 = \overset{(a)}{\frac{3 \div 2}{4 \div 4}} = \overset{(b)}{\frac{3 \div 2}{1}} =$$

$$\overset{(c)}{3 \div 2} \text{ or } 3/2 = 1\tfrac{1}{2}$$

Note: In the actual learning situation the students soon generalize that dividing the denominators always results in an answer of one. Once this generalization is made, they need only divide the numerators and may omit steps (a) and (b).

PROCEDURE

The population for this investigation was chosen from the St. Paul, Minnesota Public Elementary Schools. Twenty classes, organized on the self-contained classroom basis, were chosen from those schools having two sixth grade classes.

After the ten schools were selected, one class from each school was assigned to the control group and one class was assigned to the experimental group. This assignment was made by a coin flip. Since, within a school, the St. Paul system assigns its pupils to their classes in a random fashion, it was assumed that the groups would be of equal ability.

The summary of the analysis of intelligence data, utilizing the analysis of variance, is given in Table 1. The intelligence measure used was the Otis Intelligence Test (Beta Form) for which scores were readily available on all children. Other measures used were computation tests in addition and subtraction, multiplication, and division of fractions. These measures included examples of all levels of difficulty.

On the day prior to the initiation of the study the four fraction tests were administered. The analysis of these scores is summarized in Table 2. The

TABLE 1

ANALYSIS OF VARIANCE OF INTELLIGENCE TEST SCORES

Source of Variation	Degrees of Freedom	Mean Square	F	Hypothesis
Schools	9	249.116	26.78	.01 Reject
Sex	1	73.713	7.92	.01 Reject
Method	1	18.091	1.94	Accept
Sex × Method *	1	4.691	.50	Accept
Error	27	9.304		

* Sex × Method refers to interaction

TABLE 2

ANALYSIS OF VARIANCE OF FRACTIONS PRETEST SCORES

Source of Variation	Variation	Degrees of Freedom	Mean Square	F	Hypothesis
Addition and Subtraction	School	9	22.138	5.37	.01 Reject
	Sex	1	2.116	.51	Accept
	Method	1	6.241	1.51	Accept
	Sex × Method	1	3.969	.96	Accept
	Error	27	4 121		
Multiplication	School	9	51.665	4.56	.01 Reject
	Sex	1	65.281	5.76	.05 Accept
	Method	1	2.653	.23	Accept
	Sex × Method	1	.341	.03	Accept
	Error	27	11.330		
Division	School	9	4.548	2.19	Accept
	Sex	1	.381	.18	Accept
	Method	1	3.783	1.82	Accept
	Sex × Method	1	.551	.26	Accept
	Error	27	2.077		

scores on the initial measures were used for purposes of comparison and control of the two groups.

The experimental group was then instructed by teachers using a course of instruction prepared by the investigator. In the preparation of the course of instruction every effort was made to present the lessons in a sequence similar to the order of the material utilized by the control group. The number of practice examples assigned was equal for the two groups. Each child in the experimental group had his own personal copy of the lessons which was organized in textbook form. The control group followed the standard course of instruction which existed in the textbook utilized by the St. Paul Public Elementary Schools.[2]

The duration of the study was three weeks. During this time each teacher spent forty minutes per day on arithmetic, including seatwork. The forty minutes assigned to arithmetic was in accordance with the curriculum policies of the St. Paul Elementary Schools.

At the close of the three week period, the four fraction tests were again administered to the children. Data secured from this administration of the test were utilized for the purposes of control and comparison of the control and experimental groups. During the three weeks following the close of instruction, the experimental and control groups resumed work in their regular texts. The particular unit was a review of addition and subtraction of decimals. Children were given no instruction in or review of fraction processes during the three week period.

At the close of the three-week period the four fraction tests were administered for the third and final time. Scores obtained were used for the purposes of comparison of the experimental and control groups.

The final data used for the comparison of the two groups were obtained by taking the differences between the scores on the second and third administrations of the fraction tests. These data were utilized to make comparisons of the net gains or losses of the two groups during the three-week retention period. Unweighted means were utilized in the analysis of all data.

RESULTS AND DISCUSSION

The analysis of the intelligence test scores is presented in Table 1. The null hypotheses tested in the statistical design were:

> In reference to the intelligence test scores on the Otis Beta Intelligence Tests, there are no significant differences
>
> A. Among the ten school means
> B. Between the two sex means
> C. Between the two method means
> D. Between the method differences for the two sexes (the interaction of sex and method is equal to zero).

Table 1 indicates that in reference to intelligence test scores there were significant differences among the ten schools and between the sexes. The difference between sexes was in favor of the girls. No significant differences existed between methods or between method differences for the two sexes. Therefore, we may conclude that the experimental and control groups were equal in mental ability as assessed by the Otis Intelligence Test (Beta Form).

Analysis of the pretest scores is summarized in Table 2. The null hypotheses tested in the analysis of the pretest scores were:

> In reference to addition and subtraction, multiplication, and division of fractions prior to instruction there are no significant differences
>
> A. Among the ten school means
> B. Between the two sex means
> C. Between the two method means
> D. Between method differences for the two sexes (the interaction of sex and method is equal to zero).

From the analysis of the pretest data the conclusion can be drawn that prior to instruction the experimental and control groups showed no significant differences in their ability to add and subtract, multiply and divide fractions.

Posttest data, gathered after three weeks of instruction were analyzed and the analysis is summarized in Table 3. The null hypotheses tested in the statistical design were:

> In reference to addition and subtraction, multiplication, and division of fractions immediately after instruction there are no significant differences
>
> A. Among the ten school means
> B. Between the two sex means
> C. Between the two method means
> D. Between method differences for the two sexes (the interaction of sex and method is equal to zero).

Major conclusions can be drawn after a study of Table 3. They are (1) the experimental and control groups were significantly different in the ability to multiply fractions, the difference being in favor of the control group, and (2) the experimental and control groups were not significantly different in the ability to divide fractions.

In explanation of the first conclusion just mentioned, one can hypothesize that the inversion method, taught to the control group, utilizes multiplication of fractions as part of the process while the common denominator method, taught to the experimental group, does not involve multiplication of fractions. Consequently, there would result a difference between the two groups in the ability to multiply fractions.

Conclusion number two indicates that either of the two methods in question is equally effective in teaching division of fractions when computational performance is used as a measure for the compar-

TABLE 3

ANALYSIS OF VARIANCE OF FRACTIONS POSTTEST SCORES

Source of Variation	Variation	Degrees of Freedom	Mean Square	F	Hypothesis
Addition and Subtraction	School	9	16.973	2.31	05 Accept
	Sex	1	1.807	.24	Accept
	Method	1	1.981	.27	Accept
	Sex × Method	1	9.119	1.24	Accept
	Error	27	7.338		
Multiplication	School	9	43.095	6.47	.01 Reject
	Sex	1	32.943	4.95	.05 Accept
	Method	1	199.363	29 94	.01 Reject
	Sex × Method	1	9 701	1.45	Accept
	Error	27	6.657		
Division	School	9	89.997	6.63	.01 Reject
	Sex	1	90.601	6.68	.05 Accept
	Method	1	.001	.00	Accept
	Sex × Method	1	.576	.04	Accept
	Error	27	13.561		

TABLE 4

ANALYSIS OF VARIANCE OF THE FRACTIONS RETENTION TEST SCORES

Source of Variation	Variation	Degrees of Freedom	Mean Square	F	Hypothesis
Addition and Subtraction	School	9	24.920	4.03	.01 Reject
	Sex	1	.133	.02	Accept
	Method	1	5.853	.95	Accept
	Sex × Method	1	1.055	.17	Accept
	Error	27	6.170		
Multiplication	School	9	55.601	5.19	.01 Reject
	Sex	1	60.025	5.61	.05 Accept
	Method	1	46.225	4.32	.05 Accept
	Sex × Method	1	.256	.02	Accept
	Error	27	10.698		
Division	School	9	145.448	12.31	.01 Reject
	Sex	1	104.007	8.80	.01 Reject
	Method	1	7.141	.60	Accept
	Sex × Method	1	2.255	.19	Accept
	Error	27	11.809		

ison. However, if the common denominator method is used as the mode of instruction it would seem practical, if not necessary, to make special provision for the maintenance of skills in multiplication of fractions.

The analysis of data gathered three weeks after the close of instruction is summarized in Table 4. The null hypotheses tested in the statistical design were:

> In reference to addition and subtraction, multiplication, and division of fractions three weeks after the close of instruction there are no significant differences
>
> A. Among the ten school means
> B. Between the two sex means
> C. Between the two method means
> D. Between method differences for the two sexes (the interaction of sex and method is equal to zero).

Some very interesting information was revealed as a result of the analysis of the retention test scores. The method difference which existed in skills of multiplication of fractions immediately after instruction depreciated during the three-week retention period and was no longer significant (.01). This might lead to the conclusion that the effect of retroactive inhibition exhibited immediately after instruction had tended to become less pronounced with the passing of the three-week period during which time no instruction in fractions took place. This particular phenomenon is discussed further in the section dealing with the analysis of the data on net gain or loss during the three-week retention period.

The final data to be analyzed were the amount of gain or loss which took place during the three-week retention period. The quantitative score was derived by subtracting the score on the retention test, which was given three weeks after the close of instruction, from the score on the posttest, which was given immediately at the close of instruction.

These scores were then analyzed by the analysis of variance, the summary of which appears in Table 5. The null hypotheses tested in the statistical design were:

> In reference to gain or loss in addition and subtraction, multiplication, and division of fractions, as measured by the difference between the posttest and retention test scores, there are no significant differences
>
> A. Among the ten school means
> B. Between the two sex means
> C. Between the two method means
> D. Between method differences for the two sexes (the interaction of sex and method is equal to zero).

The analysis of the addition and subtraction test scores indicates that the amount gained or lost was (1) equal for boys and girls, and (2) equal for the experimental and control groups. Also, there was no significant interaction of sex and method.

In reference to multiplication of fractions a significant difference was found to exist between the experimental and control groups, the difference being in favor of the experimental group. In other words, the experimental group gained significantly while the control group showed no significant change.

The analysis of the scores in divisions of fractions revealed no significant differences. Therefore, we may conclude that there is no discernable difference between the two methods in relation to short-term retention.

The analysis of covariance was utilized where it seemed feasible to control pre or posttest scores and in order to substantiate the results found when the analysis of variance was employed. In each case, the analysis of covariance revealed findings which were in agreement with those obtained by the analysis of variance.

SUMMARY OF MAJOR FINDINGS

1. Girls scored significantly (.01) higher than the boys on the Otis Beta Intelligence Test.
2. The inversion method of teaching division of fractions tends to provide for significantly (.01) better results on the multiplication of fractions posttest. It would appear that the inversion method of dividing fractions tends to reinforce skill in multiplication of fractions. On the other hand, the common denominator method of dividing fractions either a) tends to confuse the children in their skill of multiplication of fractions or b) does not reinforce skills in multiplication of fractions.
3. Girls scored higher (.05) than boys on the multiplication of fractions post and retention tests. One could conclude that this difference was the result of true differences in intelligence. However, this finding is in the region of doubtful significance and further research is required before definite conclusions can be drawn.
4. Girls scored significantly (.01) higher than boys on the division of fractions retention test. During the three-week period without review, the girls retained significantly more skill than boys in the ability to divide fractions as measured by the division of fractions retention test. We could conclude that this difference was in fact due to differences in intelligence, but due to the nature of the intelligence measure used, these results must be interpreted with caution.
5. On the multiplication of fractions retention test, the experimental and control groups showed a difference (.05) in the ability to multiply fractions. This difference is in the region of doubtful significance. However, it appears that during the three-week period without review, the degree of difference in the ability to multiply fractions became small-

TABLE 5

ANALYSIS OF VARIANCE OF DIFFERENCES BETWEEN THE FRACTION POST AND RETENTION TEST SCORES

Source of Variation	Variation	Degrees of Freedom	Mean Square	F	Hypothesis
Addition and Subtraction	School	9	3.508	2.38	.05 Accept
	Sex	1	.961	.65	Accept
	Method	1	1.024	.69	Accept
	Sex × Method	1	3.969	2.70	Accept
	Error	27	1.469		Accept
Multiplication	School	9	8.976	1.72	Accept
	Sex	1	4.033	.77	Accept
	Method	1	53.593	10.30	01 Reject
	Sex × Method	1	6.805	1.30	Accept
	Error	27	5.202		
Division	School	9	15.881	1.92	Accept
	Sex	1	.463	.05	Accept
	Method	1	6.973	.84	Accept
	Sex × Method	1	.551	.06	Accept
	Error	27	8.234		

er than when measured immediately after instruction. When measured immediately after instruction the differences were significant at the .01 level of probability.

6. The experimental group gained significantly (.01) in the ability to multiply fractions. This finding supports the results stated in number five. Some of the difficulty that existed in multiplication of fractions in the experimental group immediately after instruction tended to disappear during the three-week period after instruction. Conversely, the control group tended to maintain approximately the same degree of efficiency during the three-week period.

7. In reference to the ability to divide fractions, neither method is superior to the other in terms of initial achievement or short-term retention.

8. There was no significant interaction of sex and method.

9. Neither method showed any significant effect upon the ability to add and subtract fractions.

CONCLUSIONS

Division of fractions, as defined by the test in this study, can be taught by either method with a fair degree of success in the three-week period alloted for instruction. The common denominator method of teaching division of fractions results in approximately the same degree of achievement in ability to divide fractions, as measured by the fractions test, as does the inversion method of teaching division of fractions.

In reference to interference with and strengthening of other fraction skills measured in this investigation, it appears that the common denominator method results in some difficulty for the children in their ability to multiply fractions. This effect may be a direct result of the common denominator method since this method ends as a division exercise, thereby eliminating the opportunity to strengthen multiplication of fraction skills. On the other hand, the inversion method of dividing fractions ends as a multiplication of fractions exercise, thereby strengthening skills in multiplication of fractions. However, more research is needed on this question before definite conclusions can be drawn.

Sex differences reported in this investigation seem to be due largely to the sex difference in intelligence, the girls having scored significantly higher than boys on the intelligence measure used.

EDUCATIONAL IMPLICATIONS

The fundamental implication for teachers, resulting from this investigation, is that either method of teaching division of fractions may be used. However, if the common denominator method is used, it is necessary to make separate provision for the maintenance of skills in multiplication of fractions.

A further implication for researchers who might further investigate the two methods of teaching division of fractions relates to the teacher variable. Those teachers who taught the experimental group had not heard of the common denominator method prior to this investigation. Consequently, they had to learn the method themselves as well as teaching the material to the children. Essentially then an unfair comparison may have taken place. In the experimental group the teachers were dealing with material which was unfamiliar to them while in the control group the teachers were thoroughly familiar with the material they were presenting to the children. If this difference could be removed, or controlled, the results might be substantially different.

There is also the possibility of further research which might deal with a combination of the two methods. At any rate, it would seem well worthwhile to pursue the investigation of the merits of the two methods until more definite conclusions can be made in favor of either of the methods.

FOOTNOTES

1. The above report is based on the author's unpublished Ph. D. thesis, "A Comparison of the Common Denominator and Inversion Method in Teaching Division of Fractions", submitted to the University of Minnesota faculty.

2. Clark, John R., Charlotte W. Junge and Harold Moser, Growth in Arithmetic, Revised 6 (New York: World Book Company, 1957).

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 10, July-August 1963)

Elementary School Teachers Rate Report Cards

NORMAN M. CHANSKY
University of North Carolina, Raleigh

A CHALLENGING but foreboding task for the teacher is informing a parent of the progress his child is making in school. A wide variety of judgments about a student's achievement may be held by a teacher, but she must confine her evaluations to the reporting framework adopted by the school. School policy not only determines the areas of student growth to be evaluated but the very terminology to be used in reporting on student progress.

The report card purports to channel to parents information about a child's progress in school. Yet few planned inquiries exist as to how effective this channel is. The purpose of the present study was to determine whether teachers believe that various reporting forms are equally effective in relaying functional information about the growth of elementary school children.

Procedure

Twenty-five reporting devices were selected at random from many used in the elementary schools of New York. Seventy teachers from more than 30 school districts were asked to judge these devices. Each judge had been teaching at least two years. To insure rating according to objective criteria, the familiar goals for the elementary school were stated, and the rating of each card was to be made by referring to these goal statements.

The specific directions follow:

Twenty-five pupil progress reports are presented below. You are to study the card and then evaluate it in light of the following criteria:[1] Does the reporting method yield functional information about the child's 1) Organismic development: physical education, public health, safety, mental health, and self-understandings; 2) Social development: Observance of social standards, moral and civil laws; trustworthiness, honesty, cooperativeness; 3) Geopolitical development: relation of child to community, state, nation, universe; understanding of geographic, civic, economic, and historic facets of American way of life; 4) Development in awareness of natural environment: physical, biological, chemical and ecological facets of life; 5) Aesthetic development: appreciation and expression in art, music, handicrafts, cooking; 6) Development of verbal communication: constructive use of reading, writing, listening and talking; 7) Development of non-verbal communication: gesture, number system, arithmetical operations.

You are to judge each card on a scale of eight points. If a report card contains no functional information relevant to the child's development in these seven areas, score the card zero. If it contains functional information relevant to at least four criteria, score it anywhere from one to four. If it contains functional information relative to five or more criteria, score it from five to seven. Remember, you may give a report card a score ranging from zero to seven, depending on how well the card fulfills all of the above criteria.

The ratings were collected and the mean, median and standard deviation of the ratings were tabulated for each card. Table 1 contains these statistics as well as essential features of the report cards.[2]

Reliability and Validity of Ratings

To determine whether the ratings were reliable, the set of cards was readministered to a subsample 32 teachers enrolled in the writer's course in testing. The median test was performed to determine whether the ratings changed during the interim. According to Table 2, any change in median rating may be attributed to chance.

The validity of the ratings could not be tested directly due to the absence of a criterion measure. The writer did, however, investigate two factors which may contaminate rating validity: 1) response bias and 2) an intelligence-reading factor: With respect to the former, the modal ratings on the odd- and even-numbered cards were found for each rater. Comparisons between modal ratings on each half were made. Table 3 reveals that modal ratings on each half are independent of one another, permitting the inference that a response bias explanation for the ratings is not tenable.

With respect to the latter, the Watson-Glaser Appraisal of Critical Thinking was administered to the raters. Total scores were correlated with the modal ratings for all 25 cards. The Pearson product-moment correlation was -0.05, indicating that a critical thinking factor is not associated with these ratings.

In summary, the ratings of the cards are generally consistent over a period of time, are not influenced by a responding bias of the rater, and are not related to a reading-intelligence factor. The writer infers that rating report cards according to the several goals for elementary education are reliable and resistant to some factors which would contaminate validity.

Results and Discussion

The data in Table 1 suggest that the ratings tend to cluster about a point on a continuum rather than being distributed in a spurious manner. For the majority of cards, there was little difference between mean and median ratings. Although total N is relatively small, there is, nevertheless, a semblance of symmetry to the ratings.

As a technique for rating cards, statements of the several goals for elementary education function as effective criteria. While no statement of the actual differences in information value of each report card can be made on the basis of these data, the results do indicate that teachers perceive differences in information value of report cards.

An examination of the report cards themselves reveals that at least three features distinguish high from low rated cards.

In the first place, the number of entries on the cards is a significant factor. Cards with low ratings (0-2) had a median number of 13 entries, with a range of 7-23; cards receiving a moderate rating (3, 4) had a median number of 25 entries, with a range of 16-30; cards with high ratings (5+) had a median number of 32 entries, with a range of 23-43. The more student characteristics to be evaluated on a card, the higher was its rating. Cards with many entries are apt to reflect progress in the many facets of learning inaugurated in school.

In the second place, the range of child behavior covered by the card was a significant factor. Of the low rated cards, one dealt with subject matter only; two dealt with subject matter and social adjustment; and three dealt with subject matter and work habits. Of the cards with moderate ratings, one dealt with subject matter, social adjustment and health; two dealt with subject matter and work habits; five dealt with subject matter, work habits, and social adjustment; and five dealt with subject matter, social adjustment, work habits, and health. In sharp contrast are the highly rated cards. All six dealt with subject matter, social adjustment, work habits, and health. It appears, then, that cards with restricted subject matter orientations receive low ratings. Cards, on the other hand, which stimulate the classroom teacher to reflect on the ''whole child'' receive the highest ratings.

Finally, the system used in informing parents of a child's progress in school was another important factor. Of the cards receiving low ratings, one used a strictly numerical system of grading; two used numbers and letters; one used numbers and letters; one used numbers and adjectives; one used numbers, letters, and adjectives; and one used letters and adjectives. Of the cards receiving moderate ratings, two used numbers and letters; six used letters and adjectives; five used numbers, letters, and adjectives; and one used adjectives. Of the cards with the highest ratings, four used letters and adjectives; one used numbers and adjectives; and one used adjectives only. The low rated cards, then, invited the teachers to designate achievement in terms of symbols, numerical or literal. The cards which were favored, on the other hand, stimulated the teacher to describe the child's achievement.

Descriptive statements, however, were included in the cards given low ratings. Examination of these comments, though, revealed a negative orientation toward children. Comments such as ''promotion in danger;'' ''disappointing'' and ''failing'' appeared on the report cards with low ratings. With respect to the same type of modest achievement, the following statements appeared on the highly rated cards: ''needs improvement,'' ''his best but not up to grade'' and ''below average for grade.'' Threats to the child's self-esteem appeared on the cards with low ratings; the preferred cards preserved the dignity of the child.

These factors, it is hypothesized, result in differences in the effectiveness of communication between teachers and parents. As a concise summary of a child's progress in school, some reporting techniques fall short of the mark, whereas others convey to the parent functional information.

It remains for future studies to determine whether quantifiable teacher preferences for reporting systems are related to such variables as teacher attitude toward children, teacher philos-

TABLE 1

REPORT CARD FEATURES AND RATINGS

Card	Number of Entries	Growth Areas	Evaluation Samples	Median	Rating Mean	S.D.
1	19	A, S, H, W	Yes, no — if behavior present	3.38	4.30	1.22
2	23	A, S, H, W	For grade, H, excellent U, unsatisfactory	5.33	5.30	1.84
3	16	W, A	H, excellent U, failing	3.36	3.38	1.67
4	13	A, W	E, excellent U, unsatisfactory	2.41	2.71	1.75
5	7	A, W	A, excellent, 90-100 F, failing, below 70	0.00	0.68	1.00
6	27	A, S, H, W	V, very good; also A, 90-100, excellent	4.87	4.91	1.95
7	20	A, S, H	I, improving; also 90-100, A	4.56	4.61	1.11
8	32	A, S, H, W	S, satisfactory also A, excellent	6.10	5.88	0.91
9	25	A, W, S	A, excellent F, failing	4.00	4.44	1.51
10	27	A, S, H, W	A, 91-100 D, 65-74	4.16	3.93	1.72
11	30	A, S, H, W	A, very good F, failure	5.10	5.14	2.00
12	43	A, S, H, W	letter to home	5.41	5.19	2.00
13	26	A, S, H, W	E, excellent N, needs improving	3.72	4.00	1.69
14	23	A, S, W	A, excellent, 90-100 D, failing, needs improving, 65-74	3.50	3.78	1.58
15	30	A, S, H, W	S, satisfactory U, unsatisfactory	4.50	4.70	1.79
16	41	A, S, H, W	0-100; outstanding	5.20	5.45	1.49
17	24	A, S, W	A, excellent F, failing	4.33	4.34	1.89
18	27	A, S, W	A, excellent, 90-100 F, unsatisfactory	4.30	4.33	1.81

TABLE 1 (cont.)

Card	Number of Entries	Growth Areas	Evaluation Samples	Median	Rating Mean	S.D.
19	19	A, S	0-100; Very commendable; Disappointing	2.50	2.74	1.84
20	18	A, W	A, 95-100 E, below 75	3.00	3.11	1.87
21	12	A	0-100; below 75 failing	1.40	2.17	1.67
22	16	A, W	A, 90-100 promotion in danger +, excellent	1.69	2.43	1.79
23	29	A, S, W	A, very good; check, improvement needed	4.60	5.00	1.43
24	17	A, S	A, 93-100 D, 0-74	2.60	2.84	1.62
25	36	A, S, H, W	C, commendable for grade: 1, above; 3, below	5.80	5.83	1.37

* Growth areas: A, subject matter; W, work habits; S, social-emotional learning; H, health practices

TABLE 2

TEST RETEST MEDIAN VALUES OF RATINGS OF EACH REPORT CARD AND RESULTS OF MEDIAN TESTS

Report Card	Median 1	Median 2	Grand Median	X_c^2 *
1	4	3	3	0.65
2	6	7	6	1.03
3	4	2	3	2.08
4	3	2	2	0.90
5	0	0	0	0.00
6	5	5	5	0.00
7	5	4	4	1.40
8	6	6	6	0.12
9	4	4	4	0.39
10	5	4	4	0.35
11	5	5	5	0.01
12	6	6	6	0.21
13	4	3	4	0.42
14	4	4	4	0.51
15	5	5	5	0.00
16	6	5	6	0.12
17	3	4	3	0.01
18	4	4	4	0.02
19	2	2	2	0.36
20	3	2	2	0.25
21	1	1	1	0.19
22	2	2	2	0.58
23	5	4	5	0.23
24	2	2	2	0.01
25	6	5	6	1.57

---

* 5% Probability level with one degree of freedom, $X^2 = 3.84$

ophy of learning, and teacher effectiveness.

Summary and Conclusions

A study was conducted to determine whether there were differences between report cards with respect to the information they contain. Seventy elementary school teachers judged 25 report cards in terms of the functional information they contained relevant to the commonly accepted goals for elementary education. Teachers report that cards differ with respect to the functional information they contain. Number of items on the cards, number of areas to be evaluated, and nature of the marking system were found to contrast high from low rated cards.

TABLE 3

COMPARISONS BETWEEN RATINGS ON THE SEPARATE HALVES

Mode of Responding	Observed Frequency	Chi-square
Same mode	9	
Different mode	14	
No mode	9	1.58

FOOTNOTES

1. Adapted from Kearney, N.C., Elementary School Objectives (New York: Russell Sage, 1953).
2. The report cards used in the study are available to research workers. They may be obtained from the writer.
3. In an unpublished study, Minnesota Teacher Attitude Inventory was found to correlate +.56 with preferences for cards 3, 8, 13, 15 and 21, and +.48 with preferences for cards 5, 6, 14, 19 and 25.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 10, July-August 1963)

The Outcomes of Curriculum Modifications Designed to Foster Critical Thinking[1]

NORMAN E. WALLEN,
VERNON F. HAUBRICH,[2]
and IAN E. REID
University of Utah

CRITICAL THINKING appears to be a universally accepted objective of education though we are frequently unclear as to what we mean by it and to what extent we wish to live with its consequences. As has been pointed out elsewhere (5) various definitions of critical thinking seem to encompass some or all of the following features:

1. Use of scientific methods including emphasis on evidence and the nature of hypotheses.
2. The tendency to be inquisitive, critical and analytical with respect to issues, personal behavior, etc. A derivative of this attribute is lack of susceptibility to propaganda.
3. Use of correct principles of logic.

The emphasis is on the development of that elusive philosophical ideal, the rational man.

With respect to methods of fostering critical thinking two major approaches have been advocated. The first is "progressive education." Critical thinking is presumed to be but one of the objectives which are fostered by a greater degree of self determination, flexibility of curriculum and freedom of behavior. The results of the Eight Year Study provide some support for this position. Further support of an indirect type is provided by studies which indicate that questioning and critical behaviors are less likely to occur in rigid, highly formalized situations wherein deviation is punished (2).

The second approach emphasizes the tools rather than the attitude of critical thinking while recognizing the importance of a milieu conducive to the use of the tools. Thus emphasis is placed on acquainting students with the principles of logic and experimentation and with their use. It is this approach to which this study was directed.

METHOD

The basic design of the study was as follows: It involved seven teachers of U. S. History (eleventh grade) in three Salt Lake City high schools who introduced the curriculum modifications and an additional two who served as controls. During the first year one class (selected at random) taught by each of the nine teachers was tested in the fall and again in the spring to establish the amount of gain to be expected over a year's time under the present curriculum. The tests used were the Cooperative U. S. History Test, the Watson-Glaser Test of Critical Thinking and the I. D. S. Critical Thinking Test. During the summer of 1960, the experimental teachers attended a one week workshop on the University of Utah campus under the direction of Dr. Haubrich during which time they received training in the curriculum procedures and materials presently available as well as experience in the development of new materials. During the following academic year two of their classes were again tested in the fall and spring as were those of the control teachers. During this year, the staff members worked with the teachers in the utilization and development of materials. The resulting data permitted comparisons of gains made from year to year under the same teacher and from teacher to teacher within a given year.

The statistical analysis used was analysis of covariance, which permits comparison of end of year scores – adjusted for beginning of year scores under the different treatments. Thus (in effect) the mean gain achieved by the experimental teachers during the first year – regular curriculum– is compared with the mean gain achieved under the modified curriculum. Further, the mean gain achieved by the experimental teachers using the modified curriculum is compared with the mean gain achieved by the control teachers during the same year.

CURRICULUM MODIFICATIONS

The overall plan of curriculum modification called for the teaching of a unit in "critical thinking" followed throughout the year by application to the content of the course as rather broadly defined. As an example, the students were encouraged to examine their textbook, their newspapers and their teachers for examples of fallacious logic. This approach has been extensively developed in the Illinois Curriculum Program under the direction of B. Othanel Smith and his associates. In a comprehensive application of the plan in Illinois, a total of 36 teachers and approximately 1,500 high school students in English, geometry, science and social studies classes participated. As of this writing, only a preliminary report has been published (5). It appears that the study was carefully conducted and that the students experiencing the experimental method showed greater gain on measures of critical thinking than the control group without showing impairment in mastery of course content.

Thus, the present study is, to a large extent, a replication of the Illinois study to determine whether similar results are obtained – a procedure woefully lacking in educational research. In addition the present study contains some methodological improvements, notably the use of a "base line" for guaging change which is based on the same teachers who institute the curriculum changes.

For convenience, the curricular practices may be divided into (1) materials presented during the unit on critical thinking, and (2) materials used throughout the remainder of the year.

1. Unit on critical thinking. This unit required approximately three weeks for all teachers and was conducted – at the teachers convenience – sometime during the second or third month of school.

The sequence of presentation varied from teacher to teacher but included the following topics and in this general order:

a. Definitions – abstract and concrete
b. Logical fallacies – post hoc fallacy etc.
c. Deductive principles
    Syllogisms
    If – then statements
    Validity and truth
d. Inductive principles
    The nature of evidence
    Analysis of arguments including recognition of implicit assumptions
    Reliability of sources

In addition to their notes and experiences during the workshop, the teachers were provided with copies of "Applied Logic" by Little, Wilson, and Moore and copies of "Guide to Clear Thinking" developed by the Illinois Curriculum Program. Also, it was intended that each student be provided with or have access to "A Guide to Logical Thinking" by Shanner. In one school, however, a misunderstanding resulted in these booklets not being available to all students.

As can be seen from the topics listed above, the intent was to present to these students many of the more salient developments in the areas of logic, semantics and philosophy of science but in a fashion which they would comprehend.

2. Application. Throughout the remainder of the year the teacher attempted to utilize the ideas and skills taught during the unit whenever feasible. To this end many of the exercises developed by the Illinois group were used. Also, the teachers showed considerable ingenuity and expenditure of effort in materials which they developed. Some of the flavor of the materials may be conveyed by the following illustrative exercises.

a. A statement on page 77 of the text states: "The Articles of Confederation granted considerable power to a Congress of the United States." Is this definition, explanation or opinion? What criteria are provided?

b. Analyze the argument for unfair advantages of big business on page 368 of the text. Are there irrelevancies? Fallacies? Do the reasons justify the conclusion?

c. Is there a fallacy in the following argument? Life under a strong central government in Great Britain was tyrannical. We must not allow a strong central government to develop in this country.

TESTS USED TO EVALUATE OUTCOMES

The measuring devices used to assess the outcomes of the program included the Watson-Glaser Critical Thinking Appraisal, the I. D. S. Critical Thinking Test, both constructed to assess skills in critical thinking and the Cooperative American History Test which was used to assess change in the more typical content of the course.

1. Watson-Glaser. This test was originally published in 1942 and was revised in 1956. It contains five sub-tests: inference, assumptions, deductions, interpretation and arguments. It has been used in numerous studies and is quite adequate in terms of technical considerations such as reliability, norms, etc. Ennis (3) has, however, questioned its validity on the grounds that some items are questionable and that it gives too high a score to the "chronic doubter."

2. I. D. S. Test. This test was developed in 1957 by Ennis in part as an attempt to overcome his objections to the Watson-Glaser. As such the items are on logical grounds, superior. Preliminary data suggest that it is adequate from a technical standpoint.

3. Cooperative American History Test. This test is considered to be one of the best standardized tests of the typical content of American History courses. It contains items designed to test knowledge of historical facts: understanding of cause and effect relationships, trends and developments: and ability to recognize chronological relationships, interpret historical maps and locate historical information with emphasis on political and diplomatic history. It is somewhat weak in the

TABLE 1

ANALYSIS OF COVARIANCE - EXPERIMENTAL TEACHERS ONLY*

I. D. S. Test

Source of Variance	Σx^2	Σxy	Σy^2	d. f.	Adj. Σy^2	M. S.	F	P
Between years (curricula)	4	27	189	1	159	159	8.83	$<.01$
Between Teachers	423	347	428	6	168	28	1.56	
Interaction	187	229	489	6	287	48	2.67	$<.05$
Residual	8865	5066	10245	406	7350	18		
Total	9479	5669	11351	419				

Watson-Glaser

Source of Variance	Σx^2	Σxy	Σy^2	d. f.	Adj. Σy^2	M. S.	F	P
Between years (curricula)	44	58	77	1	15	15	.32	
Between Teachers	888	732	654	5	59	12	.26	
Interaction	1219	1374	1732	5	373	75	1.59	
Residual	28042	20298	31108	348	16414	47		
Total	30193	22463	33570	359				

Cooperative U. S. History Test

Source of Variance	Σx^2	Σxy	Σy^2	d. f.	Adj. Σy^2	M. S.	F	P
Between years (curricula)	904	-104	12	1	899	899	18.65	$<.001$
Between Teachers	1530	1235	1200	6	216	36	.75	
Interaction	634	725	1083	6	291	48	1.00	
Residual	23988	21601	39027	406	19575	48		
Total	27056	23457	41322	419				

* With the exception of the F column, decimals have been omitted to simplify the tables. Cases were deleted at random so as to obtain samples of 20 each for each teacher for year 1 and 40 for each teacher for year 2. This procedure necessitated dropping the classes of one teacher from the Watson-Glaser analysis since only 12 students took both test and re-test during year 1.

TABLE 2

ANALYSIS OF COVARIANCE - EXPERIMENTAL VS. CONTROL TEACHERS - YEAR 2 ONLY

I. D. S. Test

Source of Variance	Σx^2	Σxy	Σy^2	d. f.	Adj. Σy^2	M. S.	F	P
Between Groups	124	216	375	1	154	154	7.22	$<.01$
Residual	8203	5066	11352	386	8222	21		
Total	8327	5282	11727	387				

Watson-Glaser

Source of Variance	Σx^2	Σxy	Σy^2	d. f.	Adj. Σy^2	M. S.	F	P
Between Groups	254	206	167	1	24	24	.47	--
Residual	12262	13789	35229	391	19790	51		
Total	12516	13965	35396	392	19814			

Cooperative U. S. History Test

Source of Variance	Σx^2	Σxy	Σy^2	d. f.	Adj. Σy^2	M. S.	F	P
Between Groups	226	314	435	1	49	49	1.09	
Residual	22198	20397	36224	386	17481	45		
Total	22424	20711	36659	387	17530			

TABLE 3

MEANS OF EXPERIMENTAL AND CONTROL GROUPS IN THE PRESENT STUDY AND OF OTHER COMPARISON GROUPS ON THE I. D. S. TEST

	N	Mean Fall	Mean Spring	Gain
Experimental Teachers - Regular Curriculum - Year 1	140	8.8	10.2	1.4
Experimental Teachers - Modified Curriculum - Year 2	280	9.1	11.8	2.7
Control Teachers - Regular Curriculum - Year 1	36	6.8	8.4	1.6
Control Teachers - Regular Curriculum - Year 2	53	7.5	9.0	1.5
Normative Data - High School Juniors*			9.0	
Normative Data - High School Seniors*			9.6	
College Educational Psychology Students*			12.3	
High School Students in Courses Emphasizing Critical Thinking*			12.1	

*Ennis, R. H. "Interim Report: The Development of the I. D. S. Critical Thinking Test."

TABLE 4

MEANS OF EXPERIMENTAL AND CONTROL GROUPS ON THE WATSON-GLASER TEST

	N	Mean Fall	Mean Spring	Gain
Experimental Teachers - Regular Curriculum - Year 1	120	62.3	64.9	2.6
Experimental Teachers - Modified Curriculum - Year 2	240	61.6	64.0	2.4
Control Teachers - Regular Curriculum - Year 1	30	56.8	60.0	3.2
Control Teachers - Regular Curriculum - Year 2	53	59.6	62.0	2.4

TABLE 5

MEANS OF EXPERIMENTAL AND CONTROL GROUPS ON THE COOPERATIVE U.S. HISTORY TEST (STANDARD SCORES: $\overline{X} = 50$, S = 10)

	N	Mean Fall	Mean Spring	Gain
Experimental Teachers - Regular Curriculum - Year 1	140	44.1	49.3	5.2
Experimental Teachers - Modified Curriculum - Year 2	280	41.3	49.7	8.4
Control Teachers - Regular Curriculum - Year 1	36	44.1	47.6	3.5
Control Teachers - Regular Curriculum - Year 2	51	39.5	46.9	7.4

area of contemporary affairs.

RESULTS

Results of the analysis of covariance comparing students of the experimental teachers for the two years are shown in Table 1. Table 2 shows the analysis of covariance comparing experimental and control classes for the second year only. Table 3 shows the means of the various groups as well as some additional data pertaining to the I. D. S. Test. Tables 4 and 5 show mean values for the Watson-Glaser and Cooperative U. S. History Test respectively. These data support the following interpretation:

1. *I. D. S. Test*

a) Considered as a group, students of the experimental teachers showed significantly greater gain ($p < .01$) the second year, i. e. under the modified curriculum, as compared to the previous year. The amount of the difference, when compared to available norms indicates the improvement to be of practical importance. The students under the revised curriculum, began the year with a mean score very near that typical of eleventh graders and, by the end of the year scored at a level almost up to that of a sample of unselected college students and almost as high as previously reported groups in high school classes emphasizing critical thinking. Students of these teachers but without the revised curriculum showed the amount of gain to be expected during the course of a year. Both groups begain the year with nearly identical mean scores.

b) The significant ($p < .05$) teacher by method interaction suggests that the curricular modifications are more effective with some teachers than with others.

c) When students experiencing the revised curriculum were compared with students in the regular curriculum (during the same year — different teachers), they showed significantly greater gain ($p < .01$). The gain for the students in the regular curriculum (two teachers) was almost identical for the two years.

It seems legitimate to conclude that the revised curriculum had a rather marked effect on critical thinking as measured by the I. D. S. Test.

2. *Watson-Glaser Test.*

a) The results for this test do not support the I. D. S. Test results. There is essentially no difference between the two groups of students taught by the experimental teachers in amount of gain. In both years, the gain is 2. 8. The group experiencing the modified curriculum was slightly lower on the fall testing. For the first year group, the gain is from a percentile score of 77 to 83 while for the second year group (modified curriculum) the gain is from the 74th to the 81st percentile rank based on high school norms. Grade equivalent scores are not available for this test.

b) The comparison of experimental and control groups during the second year only is consistent with the foregoing analysis in showing no significant difference between the groups.

The results for this test provide no evidence of the modified curriculum. This finding is particularly disappointing in light of the fact that the Illinois study did find a significant superiority in amount of gain shown on this test by the students in the experimental group.

3. *Cooperative U. S. History Test.*

a) Students under the modified curriculum made significantly more gain during the year than did students with the same teachers during the preceding year ($p < .001$). In both instances, the students at the end of the year scored slightly below national norms. The experimental group, however, scored considerably lower at the beginning of the year.

b) The experimental group (modified curriculum) showed more gain than the control group during the same year, but not significantly so.

c) The control teachers achieved significantly ($p < .05$) more gain the second year.

d) The gain of the experimental teachers was not significantly greater than the gain achieved by the control teachers during the second year. Because of the gain achieved by the experimental teachers we are tempted to suggest that the curricular modifications may have fostered greater interest and/or skill in dealing with the course content, hence, greater mastery. But since the gain was not significantly greater than that achieved by the control teachers during the second year, it is possible that other factors were operative, possibly that the second year students began the year with somewhat poorer background. It is clear that the modifications did not result in a decrease in the mastery of course content.

Reactions of teachers, students and parents. An additional measure of the outcomes of a plan such as this is to be found in the reactions of persons involved in it. Although no systematic attempt was made to collect such data in the present study, some information almost inevitably is present. It is recognized that impressions such as those which follow are subject to many criticisms on the grounds of selective sampling and bias of several kinds; they are nevertheless presented as valuable, though for the most part subjective, data.

1. The seven experimental teachers have all expressed considerable enthusiasm for the program as an interesting and worthwhile attempt in an important area, though some are quite skeptical as to the results achieved, particularly among the less able students. Even accounting for the expected desire to comfort the researchers and to justify their own efforts, it is our opinion that this represents an honest reaction on the part of the teachers. One bit of supportive data is that they have all indicated an intention to use at least part of the ma-

terials next year and have expressed the hope that further work of this kind will be undertaken.

The consensus seems to be that the material on fallacies and definitions was easiest to put across with the material on syllogisms the most difficult – as would be expected. As to organization of presentation some of the teachers indicated that they would do it again in essentially the same way. Others would prefer to spread the topics out during the year and introduce them as smaller units and one teacher would, in the future, not teach the material as a distinct unit but rather would attempt to incorporate it throughout the course.

2. As reported by the teachers the reaction of students was varied. Some expressed the view that it was difficult. Others wondered what it was for, i.e. "Why don't we just have history?" Our expectation was that some students would be psychologically threatened by the material; this seems to have been the case but to a lesser extent than we expected. On the other hand some became intrigued and enjoyed it. Several teachers reported students making use of the material in arguments and particularly in debate, though some of the same material frequently is presented in debate (and in psychology courses). Several incidents of carry-over to other activities were reported:

a) Letters were written to several advertisers and to a weather man requesting definition of terms. The former were not satisfactorily answered; the latter was – and in some detail.

b) As a result of a difference of opinion in class regarding a syllogism, several students wrote to a professor of philosphy at the University of Utah for clarification.

3. There appears to have been little reaction from parents. As expected, some parents feared that knowledge of history was being sacrificed for some new silliness but the teachers were able to provide an explanation which was at least in some cases, considered adequate.

We had expected some objection from parents along the lines that their children were beginning to question some of the eternal verities. That this did not happen may be attributable to parents confidence in the schools, to parental indifference or to lack of impact of our program.

SUMMARY

This report describes a two year project which introduced into three high schools a curriculum plan designed to foster critical thinking and which attempted to assess its effectiveness. The curriculum plan was patterned after a similar program developed at the University of Illinois and consisted of the presentation of a three week unit on the tools of logical analysis, semantics and scientific method at a level appropriate to eleventh graders followed by application of these tools to the content of the course in U.S. History throughout the year. The seven participating teachers were provided a workshop prior to the introduction of the unit and were provided the services of the project staff, as well as the benefits of several group discussions, during the year. Their interest and effort expended in the project was such as to leave no question but that the approach received an adequate trial.

The results of the evaluation demonstrate quite clearly that mastery of the typical content of the U.S. History course was not impaired by the curriculum modification. The effectiveness of the program in fostering critical thinking is not unequivocally demonstrated, since one of the tests to assess this change did not show any difference between experimental and control groups. The other test, however, which on logical grounds may be argued to be a better test, did show rather impressive differences in favor of students who received the revised curriculum. Further, the reactions of teachers and students, though not intensively studied, strongly support the value of the program.

FOOTNOTES

1. This research received financial support from the Utah Educational Research Council
2. V. F. Haubrich is now at Hunter College in the Bronx.

REFERENCES

1. Aschner, Mary J. "Teaching the Anatomy of Criticism," The School Review LXIV (1956) pp. 317-22.
2. Carpenter, Finley. "Educational Significance of Studies on the Relation Between Rigidity and Problem Solving," Science Education XL (1956), pp. 296-311.
3. Ennis, Robert H. "An Appraisal of the Watson-Glaser Critical Thinking Appraisal," Journal of Educational Research, LII (1958), pp. 155-58.
4. Ennis, Robert H. "Interim Report: The Development of the I. D. S. Critical Thinking Test," Personal Communication.
5. Henderson, Kenneth B. "The Teaching of Critical Thinking," Phi Delta Kappan, XXXIX (1958), pp. 280-82.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 10, July-August 1963)

Typing Performance as Related to Mental Abilities and Interests: A Preliminary Study

HORTENSE G. WHITE
Pearl High School
Nashville, Tennesseee[1]

THE PURPOSE of this report was to study typing performance as it is related to mental ability and to measured interests. The indices of typing performance used were:

1. Typing Speed
2. Total Typing Test Score
3. Final Typing Grade

The measures of typing speed and total typing test scores were obtained from the Twentieth Century Typing Test.[2] Indices of mental ability were obtained from scores on the Thurstones' Primary Mental Abilities Test[3], and indices of measured interest were obtained from scores on the Kuder Preference Vocational Record.[4] An additional measure of mental ability or, more specifically, scholastic aptitude, was obtained through the average of first-year high-school grades. These grades, filed in the central office of the school, had been recorded in percentage form.

PROCEDURE

The tests of mental ability were administered at the beginning of the fall semester of 1955-56 to the entire typing class concerned by a local university professor of tests and measurements. The typing test was administered to the typing class by the author of this paper, during the progress of the course in Typing 1-G, in the fall semester of 1955-56. The author had the assistance of the local university professor[5] of tests and measurements in the statistical analysis of the data. Pearson's product-moment correlational analysis was applied to the data.[6] The significance of the correlation coefficients was read directly from the table of required correlation.[7]

SUBJECTS

The subjects of the study were the entire typing class of Typing 1-G, taught by the author, during the first semester of 1955-56. There were 18 female and 2 male members of the class. All of the subjects were sophomores, taking the general education curriculum. Typing 1-G is an elective, first-course in typing, open to any student whose classification is sophomore or above. This particular class in typing was chosen for study since it was considered fairly typical[8] of other typing classes previously taught by the author.

RESULTS

The intercorrelations of Typing Speed, Total Typing Test Score, Final Typing Grade and First-Year Grade in High School for the subjects under study are given in Table 1.

According to the data in Table 1, First-Year Grade Average is significantly related, at the five percent level, to the Final Typing Grade, with a r .46. First-year grade average, however, is neither significantly related to the Total Typing Test Score (r .27) nor to the Typing Speed score (r .19). Grade averages based on all the courses taken during the freshman year are significantly related to grades in typing as taught by this particular teacher.

Typing Speed is significantly related to Total Typing Test Score, r .86 and to Final Typing Grade, r .77 beyond the .01 level. The correlation of .86 is spuriously high however, because Typing Speed is already included in Total Typing Test Score.

The Total Typing Test Score is significantly related to the Final Typing Grade, r . 86, beyond the . 01 level; is significantly related to Typing Speed, r . 86, beyond the . 01 level, but not significantly related to First-Year High School Grades, r . 27. The Typing Speed Score is part of the Total Typing Test Score is a part of the Final Typing Grade. Thus the r's of r . 86 in each case, are spuriously high.

Final Typing Grade, assigned by the teacher, is significantly related to Typing Speed r .77; to Total Typing Score r .86 and to First Year Grade Average r .46. The Final Typing Grade is more related to the Total Typing Score than to the Typing Speed or to the First Year Grade Average.

Table 2 gives the correlation coefficients for typing performance and aspects of mental abilities as defined by the Primary Mental Abilities Test. According to Table 2, Verbal Meaning is significantly related to Typing Speed but not significantly related to any of the other typing performance variables, i. e., Total Typing Score and Final Typing Grade. Verbal Meaning is defined by the test as the ability to understand ideas expressed in words. High ability in this factor is especially useful in such school courses as shorthand, history, science, foreign languages and English, according to the Manual.[9] Verbal meaning is needed for success in such careers as secretary, teacher, editor, scientist, librarian, executive.

Space is significantly related to Typing Speed and to Final Typing Grade. The correlation between Space and Typing Speed is . 58, significant beyond the . 01 level; the correlation between Space and Final Typing Grade is . 48, significant at the . 05 level. Space is defined as the ability to think about objects in two or three dimensions. The designer, engineer, carpenter, etc., are typical workers who need this ability. Space is helpful, according to the manual, in geometry, art, mechanical drawing, geography classes.[10]

Reasoning is significantly related to Total Typing Test Score, r .53 and to Final Typing Grade, r .52, at the . 05 level of significance. Reasoning is the ability to solve logical problems; to foresee and plan. It is the ability required of professional workers. The higher one goes in school, the more of Reasoning he needs for success.[11]

Number is significantly related, r .47, at the . 05 level with Total Typing Test Score. Number is the ability to work with figures; to handle simple quantitative problems rapidly and accurately. Salesclerks, accountants, bookkeepers should be high in Number. Number ability is useful for success in business arithmetic, accounting, bookkeeping.[12]

Word Fluency, the ability to write and talk easily, is significantly related to the Final Typing Grade, r .76, significant beyond the . 01 level. The manual indicates that writers and salesmen require this ability; classes in drama, speech, debate, journalism should help to develop this ability.

General Intelligence is significantly related to Typing Speed, r .47, and to Total Typing Score, r .52, at the five percent level of significance. General Intelligence is significantly related to Final Typing Grade, r .67, beyond the one percent level of significance.

In summarizing Table 2, it may be said that General Intelligence is a good indicator of Typing Performance as reflected in the Speed Score; in the Total Typing Score and in the teacher's Typing Grade. The teacher's grade was more significantly related to general intelligence than was Typing Speed or Total Typing Test Score. However, all three variables of typing performance were significantly related to general intelligence.

In terms of the specific factors of intelligence, Word Fluency is the best indicator of the Final Typing Grade. Space is the best indicator of Typing Speed. Reasoning and Number are equally good indicators of the Total Typing Test Score.

Table 3 shows the correlation coefficients between first-year high school average and specific factors of mental ability as defined by the Primary Mental Abilities Test. Number is the only mental factor which is significantly related to the first-year grade average, r .58, significant beyond the . 01 level. The only negative correlation is that with Word Fluency, r -. 09. Number, therefore, is a mental factor which is significantly related to academic achievement during the first year at high school, among the present sample. It may be said that the High School First-Year Average is not as related to typing performance as is General Intelligence (See Table 2). The author cannot account for the fact that Number was the only mental factor significantly related to first-year high school average, since the present sample did not contain students taking more math courses than any other courses. The 1-G Typing is not a specialized course for majors in business education but rather it is the first course in typing, open to all sophomores and above, regardless of the curriculum in which the student may be registered.

According to Table 4, the Scientific scale of interest is significantly related to Final Typing Grade, r .56, at the five percent level. The Scientific interest means that you like to discover new facts and solve problems. Doctors, chemists, nurses, engineers, etc., have high scientific interests.[13]

The Musical scale is significantly related to Total Typing Score, r .46, at the five percent level of significance. The Musical interest[14] shows that one likes going to concerts, playing instruments, singing or reading about music and musicians. Probably, also, the Musical interest scale involves a sense of rhythm which is needed in typing.

These two scales, Scientific and Musical, are the only two out of ten scales of interest which show significant correlations with typing performance.

It is interesting to note that the Clerical interest is negatively correlated with the Final Typing grade, r-.27. The Clerical[15] interest involves a liking of office work that requires precision and accuracy. Jobs such as bookkeeper, accountant, fileclerk, salesclerk, secretary, statistician fall in this area of interest.

The Scientific interest appears therefore to offer the best indicator of typing performance in the present sample. This scientific interest, to discover new fact and solve problems, seems to fit into the significant r found between Number– the ability to solve quantitative problems (Table 3) and the first-year school average.

Table 5 gives the correlation coefficients between interests and first-year grade average. It can be seen that there are no statistically significant correlation coefficients. The highest positive correlation is between Computational interest and first-year grade average, r.26. This finding agrees with information contained in Table 3 in that the Number mental factor was the only statistically significant correlation with first-year grade average. The highest negative correlation is between first-year grade average and the outdoor interest, r-.38. A finding of no statistically significant correlation coefficients is understandable in view of the immaturity of the subjects. These were first-year grades and the interest scores were obtained when most of the subjects were sophomores in high school. Probably the interests of the subjects have not become clearly outlined as yet.

SUMMARY

A. Intercorrelations among Typing Speed; Total Typing Test Score; Final Typing Grade and First-Year Grade in High School

1. First-year grade averages are significantly related to final typing grades. Good prospects for typing students may therefore be looked for on the basis of the first-year general course average.
2. Total Scores on the objective typing test are significantly related to final typing grades assigned by the teacher. Total scores on the typing test are more related to the final typing grades than are typing speed scores. Students wishing to obtain good grades in typing, therefore, should emphasize making good scores on the typing objective test.

B. Typing Performance and Mental Abilities

1. General Intelligence is significantly related to Typing Speed; Total Test Score on the typing objective test and to the Final Grade in Typing. General Intelligence is more related to the Final Typing Grade than to the other variables of typing performance. Good prospects for typing students may be looked for on the basis of scores on General Intelligence as defined by the Primary Abilities Test.
2. In terms of the specific factors of Intelligence, Word Fluency is more related to the Final Typing Grade than are the other factors of intelligence. Space is the best indicator of Typing Speed. Reasoning and Number are equally good indicators of Total Typing Score.
3. The specific factor of intelligence which is most related to the first-year grade average is Number. It appears therefore that the Number ability is emphasized and rewarded during the first year of high school. If the first year general average is not available, a source of prospects for typing students lies in the math courses taken during the freshman year.

C. Typing Performance and Interests

1. The Scientific interest scale, the ability to discover new facts and to solve problems, is the only interest scale significantly related to Final Typing Grade r.56 significant beyond the one percent level. The only interest scale significantly related to the Total Score made on the objective typing test was the Musical interest, r.46 significant at the five percent level. The Musical scale is defined as a preference for music in an appreciation sense as well as in a performance sense. A source of prospects for typing students lies in the Kuder Preference Record interests scores, specifically in the Scientific and Musical scales. The reliability of the Scientific scale around .90 is higher than for the Musical scale, around .87.
2. The Computational interest — a liking for working with numbers was found to be related to the first year grade average r.26; though highest of the correlation coefficients, it was not a statistically significant one. This finding fits in with the finding in Table 3 where the only significant correlation was between Number and First-Year Grade Average.

GENERAL RE-STATEMENT

Of the criteria against which typing performance was correlated, the Mental Abilities test is

TABLE 1

INTERCORRELATIONS OF TYPING SPEED, TOTAL TYPING SCORE, FINAL TYPING GRADE AND FIRST-YEAR GRADE IN HIGH SCHOOL

Typing Performance	Typing Speed	Total Typing Score	Final Typing Grade	First-Year High School Average
Typing Speed	—	r. 86**	r. 77**	r. 19
Total Typing Score	—	—	r. 86**	r. 27
Final Typing Grade	—	—	—	r. 46*

* Significant at the five percent level
** Significant beyond the one percent level

TABLE 2

CORRELATION OF TYPING PERFORMANCE AND MENTAL ABILITIES

Typing Performance	Verbal Meaning	Space	Reasoning	Number	Word Fluency	Total Intelligence
Typing Speed	r. 51*	r. 58**	r. 42	r. 38	r. 22	r. 47*
Total Typing Score	r. 41	r. 43	r. 53*	r. 47*	r. 32	r. 52*
Final Typing Grade	r. 36	r. 48*	r. 52*	r. 06	r. 76**	r. 67**

* Significant at the five percent level
** Significant beyond the one percent level

TABLE 3

CORRELATION OF FIRST-YEAR GRADE AVERAGE AND PRIMARY MENTAL ABILITIES

Mental Abilities	First-Year Grade Average
Verbal Meaning	r . 12
Space	r . 09
Reasoning	r . 18
Number	r . 58**
Word Fluency	r -. 09
Total Intelligence	r . 19

* Significant at the five percent level
** Significant beyond the one percent level

TABLE 4

CORRELATION OF TYPING PERFORMANCE AND KUDER PREFERENCE VOCATIONAL SCORES

Typing Performance	Out-door	Mechan-ical	Compu-tation	Scien-tific	Persua-sive	Artis-tic	Liter-ary	Musi-cal	Social Service	Cleri-cal
Typing Speed	-.38	-.16	-.01	.29	-.05	-.23	-.05	.29	-.08	.20
Total Typing Score	-.38	-.16	-.01	.29	-.05	-.23	.08	.46*	.09	.20
Final Typing Grade	-.27	-.06	.14	.56*	-.16	-.32	-.20	.06	.24	-.27

*Significant at the five percent level

more significantly related to first-year typing grades as measured by this course, than the two criteria of Interest Test and First-year High School Average. The Scientific scale of the Interest Test has the next highest relationship to Final Typing grades and First-year High School average has the lowest relationship to Final Typing grades; however, the relationsip is, in each instance, statistically significant.

FOOTNOTES

1. Pearl High School is a senior high school, with 1000 Negro pupils, with general education and vocational curricula in a city of 200,000 population.
2. Lessenberry and Crawford, Twentieth Century Typing Test, published by South-Western Publishing Co., (Cincinnati, Ohio, 1952).
3. L. L. Thurstone and Thelma Gwinn Thurstone, Primary Mental Abilities Test. Published by Science Research Associates, Inc., (57 West Grand Avenue, Chicago, Ill., 1949), Form Ages 11-17.
4. Frederick Kuder, Kuder Preference Record Vocational. Published by Science Research Associates, Inc., (57 West Grand Avenue, Chicago, Ill.), Form C H.
5. Emma W. Bragg, Ph. D. Associate Professor of Education and Psychology, Fisk University, Nashville, Tennessee.
6. Quinn McNemar, Psychological Statistics Wiley and Sons, (New York: 1949) p. 96.
7. E. F. Lindquist, Statistical Analysis in Educational Research. Houghton Mifflin Company, (Boston, Mass.) p. 212.
8. Typical in terms of sex; chronological age; first-year grade average; mental ability.
9. Primary Mental Abilities Test, op. cit.
10. Ibid.
11. Ibid.
12. Ibid.
13. Manual, Kuder Preference Record Vocational. op. cit.
14. Manual, Kuder Preference Record Vocational. op. cit.
15. Manual, Kuder Preference Record Vocational. op. cit.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 10, July-August 1963)

Comparison of Student Reaction to Similar, Concurrent Teacher-Content Instruction

SHIRLEY ULLMAN WEDEEN
Brooklyn College

ALMOST EVERY instructor some time in his teaching career has had the feeling that while he has attempted to give the same instruction to two or more sections of the same course, he finds that he may be giving two different courses.

This instructor had such an experience with two classes, one class predominantly lower sophomores and one mostly upper sophomores. Why, then, did the instructor feel certain the class response differed greatly from one group to the other.

The problem was to investigate whether the instructor is justified in believing that the two sections reacted differently to like teacher-content stimuli. At the end of the semester, the two classes were asked to fill out, anonymously, Brooklyn College's "Student Reaction to Teaching" questionnaire.

The results indicate that the instructor's impression that Class A reacted differently than Class B to like teaching was confirmed as presented in Table 1. Class A differs from Class B as presented in Table 2.

Conclusions and Discussion

Class A, the class which was eight months older and one semester in advance, perceived the course in a less favorable light than did Class B. Of particular interest is the striking difference in reaction to: a) "Attitude Toward Students," where 35 percent more of Class B felt the teacher more "sympathetic" than Class A; b) "Examinations," where 20 percent more of Class B felt the examinations were "about right" whereas 35 percent more of Class A felt the very same examinations were "confusing;" and c) "Grading" of the same "quizzes" where 25 percent fewer in Class B felt the grading of the quizzes were "too lax."

The findings suggest that two sections of the same course given by the same instructor at the same time with identical assignments and examinations can be perceived differently by different groups of students. Some areas of perception are more influenced than others.

TABLE 1

PERCENTAGE SELECTING EACH CHOICE IN EACH SECTION AND THE AMOUNTS OF DIFFERENCE FROM CLASS A TO CLASS B IN PERCENTAGE POINTS

	Class A	Class B	Diff. from A to B*
Teacher's Attitude Toward Subject			
Enthusiastic	85	90	+ 5
Bored	0	0	0
Teacher's Apparent Familiarity with Subject			
Extensive	95	95	0
Limited	0	0	0
Organization of Subject Matter			
Excellent	30	35	+ 5
Poor	10	15	+ 5
Presentation of Subject Matter			
Very clear	65	80	+15
Confusing	5	0	- 5
Encourages Students to Introduce, Follow-up, and Use Ideas			
Above Average	70	60	-10
Below Average	5	0	- 5
Willingness to Explain			
Welcomes Questions	100	100	0
Discourages Questions	0	0	0
Opportunity for Discussion			
Above Average	80	90	+10
Below Average	0	0	0
Encouragement to Engage in Thinking Process			
Above Average	35	45	+10
Below Average	10	5	- 5
Formality of Class Atmosphere			
Too Formal	0	0	0
Too Informal	5	5	0
Teacher's Awareness of Level and Comprehension of Class			
Above Average	45	60	+15
Below Average	0	0	0
Attitude Toward Students			
Sympathetic	45	80	+35
Antagonistic	0	5	+ 5
Amount and Difficulty of Work Load			
A. Reading			
Too Little	10	0	-10
About Right	90	80	-10
Too Much	0	5	+ 5
B. Written Work			
Too Little	5	0	- 5
About Right	95	90	- 5
Too Much	0	0	0

TABLE 1 (Continued)

	Class A	Class B	Diff. from A to B*
Distribution of Work Load Through Semester			
Good Spacing	85	90	+ 5
Poor Spacing	15	10	- 5
Giving Specific Directions or Assignments in Relation to Work Load			
About Right	65	75	+10
Too Little	30	15	-15
Examinations			
A. Length			
Too Much	0	5	+ 5
About Right	95	95	0
Too Little	5	0	- 5
B. Number			
Too Much	0	0	0
About Right	80	70	-10
Too Little	20	20	0
C. Scope			
Too Much	10	15	+ 5
About Right	85	70	-15
Too Little	5	5	0
D. Clarity			
About Right	45	65	+20
Confusing	45	10	-35
Grading			
A. Reports, Term Papers, etc.			
Too Lax	5	5	0
Fair	95	90	- 5
Too Severe	0	5	+ 5
B. Quizzes			
Too Lax	10	5	- 5
Fair	85	60	-25
Too Severe	0	0	0
Overall Teacher Contribution			
More than Expected	45	40	- 5
Up to Expectation	40	35	- 5
Less than Expected	15	15	0

Note: The percents are based on the total number of students in Class A (32) and in Class B (28). They do not add up to 100 percent when students failed to answer questions.

* + or - refers to percentage difference between the classes when + means more for Class B and - means less for Class B in terms of the question posed. Reference to "favorable" versus "unfavorable" must be made in terms of each particular question.

TABLE 2

COMPARATIVE MEANS BETWEEN CLASS A AND CLASS B FOR CHRONOLOGICAL AGE, HIGH SCHOOL AVERAGE, SCHOLASTIC APTITUDE TEST, INDEX, NUMBER OF CREDITS COMPLETED, COURSE GRADE AND COLLEGE LEVEL

	Class A	Class B	Diff. from A to B
Chronological Age	19 yrs. 7 mos.	18 yrs. 11 mos.	8 mos.
High School Average	87	87	0
S. A. T.	818	793	-25
Number of Credits	68	52	-16
Index	+32	+19	-13
College Level	U.S.	L.S.	1 semester
Course Grade	2.9	3.0	+0.1

Notes:

1. Chronological age as expressed in years and months.
2. High School Average - the composite numerical academic average over a period of eight high school terms.
3. S. A. T. - the raw score obtained on the Scholastic Aptitude Test.
4. Number of Credits - the cumulative number of course credits completed to date.
5. Index - Brooklyn College scholastic index based on a point system where F = -2, D = -1, C = 0, B = 1, A = 2, for each college credit of work.
6. College Level - expressed as U.S. for upper sophomores, L.S. for lower sophomores.
7. Course Grade - grade earned for the course under investigation based on a system where F = 0, D = 1, C = 2, B = 3, A = 4, for each college credit of work.

An Analysis of Remedial Activities Used by Elementary Teachers in Coping with Classroom Behavior Problems

DONALD L. BARNES
Ball State Teachers College
Muncie, Indiana

MOST EDUCATIONAL psychologists seem to subscribe to the theory that behavior is caused and that successful modification of behavior patterns is closely related to the nature of the remedial activities used. Most frequently it is suggested that therapeutic or remedial action should be directed towards the alleviation or removal of underlying causes.

Problem

The purpose of this study was to discover: (1) the nature of the assumptions expressed by elementary teachers in their attempts to discover factors underlying behavior patterns exhibited in their classrooms; and (2) the remedial activities inaugurated by these teachers in the attempted alleviation of these behavior problems.

Procedure

Classroom problems were identified through a series of interviews with 71 elementary teachers throughout the state of Colorado. Care was taken to secure a representative sampling from various grade levels and from a variety of socio-economic areas. A list of 50 problems was built empirically. Each teacher was asked to describe behavior problems she had experienced in her classroom. If the interviewee described a problem not already identified, it was added to the cumulative list; if the problem had been described previously by another interviewee, an additional tabulation was made. Each participant in the study was then asked to enumerate the factors which he felt contributed to the behavior problem exhibited in his classroom and describe what action, if any, he took to alleviate the problem.

Results and Discussion

The teachers interviewed listed an average of 14. 5 problems. Because of the similarity of problems identified by the interviewees, no new problems were listed among the 726 responses recorded in the last fifty interviews. The problems represented many different areas of adjustment - differences in attention spans, domination, timidity, perseverance, attention-gaining devices and emotional disturbances of various kinds.

Factors believed to be contributory to these behavior problems were grouped into forty-six categories and frequencies were tabulated. The null hypothesis was formulated that there would be no significant differences in the frequency with which “factors” were listed. To test this hypothesis of equal frequencies, the chi-square, the average frequency of all “factors” enumerated was used as the expected frequency. The observed frequencies were then compared with the expected frequencies.

Of the forty-six categories, ten with highest frequencies were found to be significant at the five percent level of significance as shown in Table 1.

A further analysis of the responses given by these teachers revealed that they could be further grouped into four areas of the child's environment. Comparisons are shown in Table 2.

Analysis of the 964 responses also revealed two rather discouraging facts:

1. Only eight percent of the respondents related the “behavior problems” they enumerated to the basic needs of children (in the process of identifying what they believed were “factors” underlying these behavior patterns).
2. The respondents recorded an average of

TABLE 1

TEACHER RESPONSES REGARDING FACTORS CONTRIBUTING TO CLASSROOM BEHAVIOR PROBLEMS

Categories Enumerated	Number of Responses	Chi-square Significant at 5% Level
1. Do not know the factors contributing to the problem.	136	537.46
2. Insufficient ability or judgment on the part of the teacher.	88	175.72
3. Differences in interests among children.	81	139.43
4. Desire for attention on the part of the children.	67	79.44
5. Differences in family values.	65	72.33
6. Insufficient interest on the part of parents.	61	58.86
7. Limited intelligence of children.	58	49.72
8. Children's poor foundation in school subjects.	47	22.80
9. Children's limited home background experiences.	40	11.04
10. Differences in children's ability to concentrate.	35	5.26

TABLE 2

RELATIONSHIP OF "FACTORS" TO AREAS IN CHILD'S ENVIRONMENT

Areas with Which "Factors" Were Associated	Number of Responses	Chi-square Significant at 5% Level
1. "Factors" associated with the child's home environment.	320	34.32
2. "Factors" not exclusively associated with home, school, or heredity.	287	11.60
3. "Factors" associated with the school environment.	208	6.09
4. "Factors" associated with heredity.	149	46.98

TABLE 3

MOST FREQUENTLY USED METHODS OF HANDLING CLASSROOM BEHAVIOR PROBLEMS

Teachers' Reactions to Classroom Behavior Problems	Number of Responses	Chi-square Significant at 5% Level
1. Took no action.	163	688.67
2. Provided special activities (enrichment activities, special assistance, diagnostic tests, etc.)	136	441.73
3. Reasoned with one or more children.	104	219.78
4. Individualized or modified the activities provided for the children.	74	81.34

1.04 responses each time they had an opportunity to list factors contributing to behavior problems. This indicates that the interviewees had a strong tendency to associate problems with a single factor rather than a complex of interrelated factors.

In attempting to cope with the fifty classroom behavior problems listed, the Colorado teachers utilized 39 different kinds of remedial activities. There were striking differences in the ways various behavior problems were handled. An average of 7.14 different remedial actions were recorded for each of the 50 problems.

The null hypothesis was formulated that there would be no significant differences in the frequency with which the various remedial activities were utilized. To test the hypothesis of equal frequencies, the chi-square test of significance was used. In computing the chi-squares, the average frequency of all remedial actions enumerated was used as the expected frequency. The observed frequencies were then compared with the expected frequencies. Seven methods of reacting to classroom behavior problems were found to have frequencies significant at the five percent level as shown in Table 3.

Of the 39 remedial activities identified, it was found that 22 were of a positive nature, nine were of a negative nature, and eight were neither distinctly positive or negative. Removing obstacles which hindered the proper functioning of the children was used more often among the positive actions than reasoning, discussion and example. The negative activities related to the withdrawal of privileges or the use of penalties. Punitive measures accounted for approximately 7.9 percent of all actions taken.

A further analysis of the data revealed:

1. The Colorado teachers interviewed seldom inaugurated more than one remedial activity for each problem they identified. (An average of 1.01 responses was recorded each time an opportunity to list a remedial activity was presented.)
2. Approximately 14.4 percent of all responses indicated no action on the part of the respondents.
3. When remedial actions were compared with the "factors" contributing to the problems as identified by the teachers, it was estimated that approximately 25.6 percent of the "factors" or their psychological concomitants were appropriately treated. (It appeared that the problems of physically handicapped children, timid children and children with low IQ's were more effectively handled than other problems.)
4. Parental aid or cooperation was sought only 4.4 percent of the time.
5. The problems enumerated by these respondents generally centered around only a few members of the class.

Summary and Conclusions

The purpose of this study was to collect and analyze elementary teachers' responses to two questions:

1. What "factors" do you believe contribute to the behavior problems you have experienced in your classroom?
2. What action, if any, did you take to alleviate the problems?

Fifty classroom behavior problems were listed by a representative sampling of 71 Colorado elementary teachers. Respondents listed an average of 14.5 problems.

When asked to identify factors underlying these classroom problems, this group of interviewees frequently indicated that they were unable to do so. Differences in children's background experiences were cited most frequently. Only eight percent related the "behavior problems" they enumerated to the basic needs of children, and single factors rather than a complex of inter-related factors were usually cited.

In attempting to alleviate the 50 classroom behavior problems, the Colorado teachers utilized 39 different kinds of remedial activities. Most frequently, they took no action at all. Remedial activities were seldom related to the "factors" which supposedly caused the undesirable behavior patterns. More positive remedial activities were used than negative approaches.

If these Colorado elementary teachers are representative of others throughout the United States, it seems clear that much more needs to be done to help elementary teachers study and interpret behavior problems of elementary age children.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 10, July-August 1963)

A Comparison of the Performance of Pre-Graded Students with Grade-Motivated Students

T. L. HAWK and L. M DeRIDDER
University of Tennessee

EDUCATION JOURNALS contain many articles criticizing the use of grades at all levels of the education system. Redl and Wattenberg,* for example, suggest that heavy concentration on grades may destroy needed motivation and that the instructor who counts on the threat of poor grades to keep a class interested slights child needs. Research reports which give any data about the performance of students when grade threat is removed are, however, unavailable. The present investigation was designed to ascertain whether students who had been assigned a final grade for the course at the beginning of the quarter would differ significantly in achievement from students whose grades were to be determined by performance on tests and a term project during the quarter. Stated in null form, the hypothesis tested was that no significant differences in achievement would be exhibited by the two groups.

Procedure

The subjects comprised the 118 students enrolled in four sections of the beginning course in Educational Psychology taught by two instructors at The University of Tennessee. The course is required of those students who plan to become certified to teach. The majority of subjects were sophomore and junior education majors but other colleges were represented with a few students who planned to become teachers.

1. Two instructors taught two sections each, one being assigned a 9:00 and an 11:00 o' clock section, the other a 10:00 and a 1:00 o' clock section. Each section met three times each week for a 50-minute period, Monday, Wednesday and Friday.

2. Early in the spring quarter, a College of Education faculty member was asked to select one of the sections taught by each instructor as the experimental group and the other section of each instructor as the control group. Prior to the first assigned test, he visited each section by prearrangement and informed the students selected as the experimental group that an experiment which involved them was in process and that each person enrolled would receive, at the end of the quarter, a grade predetermined on the basis of the point-hour average now held by each. The grades to be assigned would be determined according to the following plan:

.00 - .99, D; 1.00 - 1.99, C
2.00 - 2.99, B; 3:00 - 3.99, A

The subjects were also asked to refrain from discussing the study with their instructors, as the instructors were not to know which of their sections had been pre-graded. He informed the control groups that they were involved in an experiment but gave them no information about it. He was careful to remain an equivalent time in each group in order that no clue of his choices might be suggested to the instructor. At no time during the quarter did either instructor obtain directly or indirectly, either from the students or from this faculty member, any indication which sections had been identified as the control and experimental groups.

3. Each section was taught by the instructors in an equivalent manner, using the

---

* Redl, F., and Wattenberg, W. Mental Hygiene in Teaching (New York: Harcourt Brace and Company, 1959), page 399.

TABLE I

COMPARISON OF EXPERIMENTAL AND CONTROL GROUPS BY SEX AND GRADE POINT AVERAGE

	Section	N	Male	%	Female	%	Range of g.p.a.	Mean of g.p.a.
Experimental	9:00	34	7	.20	27	.80	1.35-3.54	2.415
	10:00	28	4	.14	24	.86	.98-3.74	2.462
Totals		62	11	.19	51	.81	1.35-3.74	2.435

TABLE II

COMPARISON OF RESULTS OF FOUR TESTS AND THE CASE STUDY BETWEEN EXPERIMENTAL AND CONTROL GROUPS

	No. Items	Experimental	Control	diff.	df	t value	Levels of Significance
Test I	60	41.39	43.86	2.47	116	2.43	.05
Test II	60	39.71	42.88	3.17	116	2.58	.05
Test III	100	64.31	72.77	8.46	116	4.95	.01
Case Study		5.81	6.35	.54	116	1.70	Insignificant

TABLE III

GRADES EARNED AND GRADES ASSIGNED TO SUBJECTS IN THE EXPERIMENTAL GROUP

Number	Assigned	Who Earned A	B	C	D	F
A	11	1	3	5	1	1
B	40		6	27	6	1
C	10		1	2	4	3
D	1				1	
F	0					
Total Grades Earned by Experimental Group		1	10	34	12	5
Total Grades Earned by Control Group		5	19	27	5	0

same lectures, discussion procedures, audio-visual aids, and child-study assignments.

4. Each student's progress was evaluated by identical procedures: a case study of a child to which the student had been assigned, and two 60-item and one 100- (final examination) multiple choice tests based upon the text.

5. During the last class period of the quarter, the faculty member who first chose and informed the control and experimental groups of these studies revisited all sections. Again he stayed an equivalent amount of time in order to prevent any clues being given the instructors. With the control groups he questioned them about their feelings toward the value of the experiment and the extent of their efforts in the course.

6. The g. p. a.'s for the control and experimental groups were so similar that no difference of any statistical significance was noted. Eighty-one percent of the students in the experimental group were women, while seventy-three percent of the control were women. No significant differences in these two proportions were found. The sex of the subjects, previous grade point averages, and time of class meetings are presented in Table I.

Discussion of Results

Mean scores for the control and experimental groups on the three tests and the case study are reported in Table II. Mean scores for the two groups taught by each instructor were tested for possible significant differences, but all differences between instructors were so slight as to be insignificant. Differences between the mean scores for the experimental and control groups on Tests I and II are both significant at the .05 level, while the differences for Test III are significant at the .01 level of confidence.

The case studies were graded by each instructor on a ten-point scale. The mean number of points, shown in Table II, for the control group was 6.35 while the mean for the experimental group was 5.81. The difference of .54 was not great enough to be statistically significant.

Grades assigned and grades earned by the 62 subjects in the experimental group are reported in Table III. A comparisom is shown also between the grades earned by the experimental and control groups. Thirty-four members of the experimental groups earned a grade one letter lower than the one assigned; fourteen, two letters lower; two, three letters lower; and one person earned a grade four letters lower than the assigned grade. Grades earned by the control group are well above those earned by the experimental group.

The findings indicate the superiority of the students in the control group in performance. One may conclude that, in the absence of motivation for a grade, the majority of the students did not exert their usual degree of effort. The staff member who visited each section at the end of the quarter found considerable evidence of a lack of motivation. For instance, in the 9:00 o'clock section he was informed by eight students that they did not believe grades had been or would be assigned according to the announced plan, while another eight said that since their grade was assigned already, they did no work in the course. In the 10:00 o'clock section he was told by 18 that they had worked less than they generally would have. They mentioned that more time was used in carrying out the case-study assignment than any other aspect of the course. Eight of the group indicated that they had not read the text at all.

It is apparent that for this group, the absence of pressure to work for a grade acted as a depressant on work output when the usual class procedures were followed. Such findings raise questions about the validity of the arguments of many educators that grades destroy motivation. The question remains, however, whether a modification of class procedures involving group planning and individual encouragement might have changed the results.

Summary

Achievement of students whose grades for a course were determined early in the quarter, according to previous grade point averages, was compared with the achievement of students whose grades were to be determined by performance on tests and a term project during the quarter. Performance of students in the pre-graded groups was found to be significantly lower on two tests and a final examination but not on a term project. The conclusion is warranted that, for this group, removal of pressure for a grade acted as a depressant on work output.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 10, July-August 1963)

The "Successful" Child: Some Implications of Teacher Stereotyping

PHILIP LAMBERT
University of Wisconsin

FOR YEARS TEACHERS and administrators have realized that some children do less well at school than might have been expected in terms of their innate intelligence and ability, while others do better than expected. Yet, although considerable attention has been devoted to children who, for a wide variety of reasons, are unsuccessful at school, almost no attention has been paid to the case of the child who does extremely well; in other words, the successful child.

A few studies have dealt with the "normal" or "average" child, for example Barker (1), Dessenberger (2), Harris (3), and Millard (4). In 1949 Volberding (5) classified 66 eleven-year-old children as successful or unsuccessful on the basis of 2) academic record, b) accomplishment quotient, and c) popularity rating by peers. No attempt was made to define the word "successful" in any other way or to study stereotypes of "success" in terms of different ages, teachers, sex, etc.

The present study was designed to discover an experimental definition of success from the viewpoint of the teacher, and to validate this definition in terms of actual children selected by their teachers as being "successful". The following questions were asked: 1) What attributes are generally considered by teachers to accompany success, and which of these attributes are felt to be the most important? 2) What are the children chosen as "successful" like in terms of sex, academic achievement, and mental ability? and 3) Are the children chosen by teachers as successful also felt to have the attributes which are thought to accompany success?

PROCEDURE

Subjects for this study were the fourteen classroom teachers (kindergarten through sixth grade) at the University Elementary School, University of California at Los Angeles. Data were collected from them on 354 children ranging in age from 5 to 12 years. The test instrument used was the Gough Adjective Check List, which has been employed extensively at the Institute of Personality Assessment by Harrison Gough in connection with his work on personality variables in adult individuals. All material was submitted to the subjects in written form.

In Step I, the subjects were given a copy of the Gough Adjective Check List, with the following instructions:

> Here is a list of 284 adjectives. Please check those adjectives which in your opinion characterize children who are successful in elementary school situations. Of those adjectives checked, please rank the ten most important adjectives in order of importance. Place a "1" in front of the most important adjective, and a "10" in front of the least important adjective of the ten most important adjectives.

Step II was carried out two weeks later: A copy of his or her class list was given to each teacher with the following instructions:

> Here is your class list. Please select the five children you consider most successful, in order. (No. 1 the most successful child in your class, No. 2 the next least successful, etc.)
>
> Now write down the names of the five children you consider to be the least successful, in order. (No. 1 the least successful child in your class, No. 2 the next least successful, etc.

In Step III, which was carried out two weeks after Step II, the teachers were given a copy of the Adjective Check List for each child they had selected in Step II, with the following instructions:

> Here are copies of the Adjective Check List for ten children in your class. In each case, please check the adjectives which you think best describe the child in question.

DISCUSSION

Of the 354 children whom the teachers were asked to judge, 175 were boys and 179 girls. Yet, the group of 70 children chosen as "successful" was made up of 42 girls and 28 boys, while the group of 70 "unsuccessful" children was composed of 29 girls and 41 boys. Apparently, many more girls than boys are seen by their teachers as being successful in school situations. This agrees with the findings of Volberding (5). From the point of view of purely academic success, too, girls tend to do better than boys. Data on each child's rank order in his or her class were available for grades 2 and above in this study. These figures were based upon total reading scores on the California Achievement Test. Of the 45 children who were among the first five in their class, 30 were girls and 15 were boys.

Interestingly enough, only 40% of the children selected by their teachers as "successful" were in the top quintile of their class from the point of view of academic standing; that is, among the approximately 5 top-ranking children in their class (average class enrollment was 25.5). See Table 1.

Four of the 45 children selected as unsuccessful also appeared in the top quintile of their class from the viewpoint of academic achievement. See Table 2.

The children selected by their teachers as "successful" or "unsuccessful" were also considered in terms of IQ. The mean score on the Revised Stanford-Binet for children at the University Elementary School is 122. (This figure is based upon several random samplings of over 75% of the school population from kindergarten through sixth grade – it has remained approximately constant over the last ten years.)

IQ scores were available for all of the children enrolled at the UES at the time of this study, with the exception of a few who had moved away from the area before they could be tested. The 41 "successful" children in groups 2 through 6 for whom data were available had a mean IQ of 127; the 43 "unsuccessful" children in these grades for whom data were available had a mean IQ of 117. The mean IQ of the top-ranking children in each class (41 of the total of 45) was 132.

Some interesting differences were brought out by comparing the nine adjectives most often chosen by the teachers in Step I, and those adjectives most frequently used to describe the actual successful or unsuccessful child in Step III.

Step I

Adjective	No. of Times Chosen
friendly	14
imaginative	14
alert	13
capable	13
confident	13
enthusiastic	13
active	12
clear-thinking	12
curious	12

Step III - Successful Children

Adjective	No. of Children
friendly	61
capable	59
intelligent	56
cheerful	53
active	51
confident	49
independent	49
alert	48
interest wide	48

In general, the ideal and the actual successful children were described in the same terms. However, in describing the ideal successful child more teachers chose words pertaining to mental abilities ("curious, imaginative, clear-thinking"), while in describing the actual children whom they saw as successful, more chose words relating to classroom attitude ("cheerful, independent"). Apparently, good behavior in class is more important than teachers think in determining whether a child will be considered successful.

These differences showed up even more clearly when the weighted score for the adjectives ranked by the teachers in Step I was considered.

Step I - Weighted Score

Adjective	Score	Step III - No. of Successful Children
curious	54	35
imaginative	47	45
resourceful	43	30
confident	34	49
self-confident	32	47
sincere	29	41
intelligent	28	56
enthusiastic	25	44
alert	24	48

What might be called "inner-directed" qualities were prominent in the teachers' picture of what an ideal successful child would be like, while "other-directed" qualities were more common among the real children who were listed as "successful".

Descriptions of the actual unsuccessful children were far less consistent than was the case for the successful children; in other words, while all successful children appear to have many traits in common, unsuccessful children are seen as doing badly in school for a wide variety of reasons. Nearly half of the unsuccessful children were described as "friendly," and a large number were seen as "intelligent."

Step III - Unsuccessful Children

Adjective	No. of Children
anxious	36

TABLE 1

NUMBER AND PERCENT OF "SUCCESSFUL" CHILDREN APPEARING IN EACH QUINTILE OF THEIR CLASS

Quintile	Number	Percent
1	18	40.00
2	13	28.89
3	9	20.00
4	4	8.89
5	1	2.23
Total	45	100.00

friendly	33
sensitive	29
immature	27
intelligent	27
excitable	26
tense	25
active	25
absent-minded	24

These generally low scores may also be partly explained by the fact that all the teachers selected fewer adjectives with which to describe their unsuccessful children.

The attempt to assess the classroom teacher's influence over the level of achievement of the individual student raises a number of interesting questions. Does the elementary school teacher tend somehow to discriminate against boys? What sorts of reinforcement (teacher approval) are given to children considered by the teacher to be successful? Do the results vary according to the intelligence level of the student population? The present study accepts and reinforces the established hypotheses that the achievement of elementary school children is often at variance with their innate intelligence and that sex is often a factor in student achievement. Also, since the experiment is highly teacher oriented, it indicates the validity of further hypotheses. Teachers hold a mental image of the successful child. The qualities of the ideal successful child tend to be inner-directed and intellectual. Teachers are also able to rate actual students in terms of success. The actual successful child (from the teacher's viewpoint) is other-directed and the type who behaves well in class. The tendency of teachers to maintain an ideal and an actual stereotype of the successful child leads to further questions for the researcher. To what extent, for example, is a student's success determined by his ability to intuit and acquiesce with the teacher's concept of success? Also, to what degree do preconceived notions of success limit the teacher's ability to understand a child's struggle to attain classroom success?

TABLE 2

NUMBER AND PERCENT OF "UNSUCCESSFUL" CHILDREN APPEARING IN EACH QUINTILE OF THEIR CLASS

Quintile	Number	Percent
1	4	8.89
2	5	11.11
3	7	15.56
4	8	17.78
5	21	46.67
Total	45	100.00

REFERENCES

1. Barker, R. G., and Wright, H. F. Midwest and Its Children, (Evanston: Row, Peterson,1955)
2. Dessenberger, V. H. W. Comparative Study of Bright, Average and Dull Students and Their Unselected Siblings. Ph. D. thesis (Temple University, 1951), unpublished.
3. Harris, A. J. "What is a 'Normal' Child?" Journal of Teacher Education, (1952) III, pp. 58-61.
4. Millard, C. V. School and Child: A Case History (Michigan State College Press, 1945) 221 p.
5. Volberding, E. "Characteristics of Successful and Unsuccessful 11-year-old Pupils", Elementary School Journal (1949) pp. 405-10.

THE JOURNAL OF EDUCATIONAL RESEARCH
(Volume 56, Number 10, July-August 1963)

Sixth-Grade Children Evaluate Twelve Cases of Malicious Mischief

ALVIN E. ROBINSON
University of California, Santa Barbara

THE PROBLEM was to evaluate the value judgments of sixth-grade students in reacting to court cases involving malicious mischief, and to compare their decisions and reasons for these decisions with the actions of the courts in handling these cases.

PROCEDURE

1. Twelve cases were selected from 150 reported court cases involving malicious mischief for presentation to 2,613 sixth-grade students who were asked to select one of four choices of action and one of four possible reasons for this action.
2. A pilot study was conducted in Porterville, California and resulted in the rewording of some statements and led to the request that teachers read the cases and answers to the classes.
3. The sample group was drawn from the United States population, using random selection techniques, in such a manner that they were representative of the general population (within one percent) as to geographic location and community size.
4. The responses were scored as correct or incorrect as they agreed or disagreed with court decisions.
5. Rank order correlation of the first 500 responses returned with the second 500, and the first 1,000 with the second 1,000 was used to ascertain the adequacy of the sample population to represent the general population of the sixth-grade children.
6. The chi-square test was used to test the significance of the difference in the responses among the four geographic regions, the four categories of community size, and between boys and girls.

RESULTS

1. The choices made by the students are presented in Table I.
2. The children chose correctly (agreed with the court decision) in cases 2 and 11. They agreed with the reason given by the court in case 1.
3. The children reported great confidence in the law enforcement agencies, and were hesitant in the face of violence, particularly toward animals.
4. Rank-order correlations indicate the sample population as more than adequate to insure reliability of the response and is representative of the sixth-grade students in the general population of the United States.
5. There were no significant differences among the students in different geographic areas, among students in different size communities, or between boys and girls.

The large disparity between the court's interpretation of the right way of acting in these cases and the children's interpretations indicates children are immature in their value judgments about such matters. While this is to be expected, it does pose a question concerning the school's responsibility to help children develop the value judgments deemed correct by society in making judgments about such matters. Human relationships are at the core of each case, and therefore are important in the social sciences.

TABLE 1

RESPONSES OF 2,613 SIXTH-GRADE CHILDREN TO THE TWELVE SITUATIONS OF THE OPINION INVENTORY, "GETTING ALONG WITH PEOPLE", EXPRESSED AS PERCENTS

	First		Second		Third		Fourth	
	Choice	Reason	Choice	Reason	Choice	Reason	Choice	Reason
Case 1	4.5*	24.8	13.8	8.2	65.2	26.1	16.3	40.7*
Case 2	3.4	4.5	38.8	7.7	3.5	60.0	54.1*	27.5*
Case 3	12.5	9.5	16.2*	52.7	13.3	18.0	57.8	19.6*
Case 4	50.5	10.6	36.6*	5.0	8.8	32.0*	3.8	52.3
Case 5	7.5*	4.5	15.5	54.5	70.8	25.4	6.0	15.4*
Case 6	12.5	27.2	21.1	29.3	59.3	39.1	6.9*	4.2*
Case 7	47.3	32.5*	27.9	33.7	17.4*	7.5	7.2	26.1
Case 8	42.8	5.4*	6.3*	9.0	14.3	52.9	36.4	32.5
Case 9	4.8*	8.6*	16.5	13.6	41.9	24.7	36.6	53.0
Case 10	1.7	13.7	10.6	14.1*	18.2*	13.1	69.3	58.8
Case 11	27.4	39.4	20.9	8.3	39.0*	23.2*	12.4	28.9
Case 12	41.7	16.7*	19.1*	9.9	28.5	34.0	10.5	39.3

* court decisions

JeR Field News

ARE MANY CONVENTIONAL STATISTICS TECHNIQUES SUPERFLUOUS OR OBSOLETE? In view of the elaboration of a linear regression model by Drs. Robert A. Bottenberg and Joe H. Ward, Jr. of the 6570th Personnel Research Laboratory, this writer believes the answer to be a categorical "yes". Although the principles may have been recognized and put in written form by some theoretical statisticians, Drs. Bottenberg and Ward present the most sensible exposition of problem formulation and statistical analysis of data that this writer has yet seen. Their book heralds the opening of many new research doors, not only for the more sophisticated, but also for the relatively inexperienced or beginning investigator.

Among the statistical tools encompassed by the Bottenberg and Ward approach are the t-test, simple and complex designs of analysis of variance, analysis of covariance, the Johnson-Neyman technique, and curvilinear multiple regression.

Furthermore, this approach overcomes or compensates for many limitations which restrict or prohibit studies employing conventional design. Among these are: disproportionate cell frequencies in complex analyses of variance, analysis of treatment effects when concomitant variables are influenced by treatments, and detection of discontinuities in criteria.

Whereas, the complexity and limitations of conventional methods often may deter many investigators from their use, the problem formulation and analysis approach of Bottenberg and Ward provide simple procedures for complex studies.

This writer encourages anyone who is even slightly interested in better understanding on how to formulate a research problem to obtain a copy of this publication for his personal or school library.

The forthcoming book to be available in June 1963, will be entitled "Applied Multiple Linear Regression", PRL-TDR-63-6, by Robert Bottenberg and Joe H. Ward, Jr., 6570th Personnel Research Laboratory, Box 1557, Lackland A. F. B., Texas.

THE FULBRIGHT-HAYS ACT OF 1961 will make available more than 800 graduate fellowships for study and research abroad during the academic year of 1964-65.

Three types of grants are available: (1) U. S. government full grants which provide round-trip transportation, tuition, books, and maintenance. This type of grant may be used in any one of 36 different countries. Countries in all parts of the world are included in this group. (2) Joint U. S.-other government grants are offered cooperatively by the U. S. government and a foreign government. The U. S. Government provides funds for travel and the foreign government provides for tuition and maintenance costs. These awards are available for study in 18 countries, primarily in Central and South America. (3) Travel-only grants are intended to supplement maintenance and tuition scholarships received from other sources. These awards are available for study in eight countries.

General eligibility requirements for these grants are: U. S. citizenship, a bachelor's degree and language proficiency necessary to carry out the proposed study project.

College students should consult their camplus Fulbright Program Adviser regarding application. Others may write to the Counseling Division, Institute of

International Education, 800 Second Avenue, New York 17, New York. All requests for application forms must be postmarked by October 15, 1963.

MORE THAN 100 LIBRARIES HAVE REPORTED ON THEIR EXPERIENCES in operating either a private library in the school or a combined public and school library in community centers. In spite of possible economical advantages, the majority of the reporting librarians concluded that these combinations serve neither students nor adults well.

The report shows that use by adults increase when the public library is moved from combined to separate quarters. Three reasons are cited for this phenomenon. Adults are reluctant to use the public library facilities located in a school. Book selections tend to be made in terms of student requirements. School regulations, rather than adult needs, usually determine the hours of service.

LEHIGH UNIVERSITY HAS COMPLETELY REORGANIZED ITS APPROACH to teacher preparation. The new program stresses a broad liberal education at the undergraduate level and utilizes an internship at the graduate level. The University will have completely phased out its present program for undergraduate majors in education by the spring of 1965.

The new program requires a period of two summers and the intervening academic year beyond the bachelor's degree. One of the semesters during the academic year is to be devoted to a full-time supervised teaching internship in a nearby cooperating school. Interns will be paid the salary of a beginning teacher during the period of internship.

Satisfactory completion of this program leads to the Master of Education for prospective elementary school teachers and a Master of Arts degree for those students intending to teach at the secondary level.

PROJECT ENGLISH, UNDER THE AUSPICES OF THE U.S. OFFICE OF EDUCATION, is supporting several areas of activity related to improvement of the teaching of English.

Contracts for the development of Curriculum Centers have been signed with Carnegie Institute of Technology, Northwestern University, University of Nebraska, Hunter College, University of Oregon, and the University of Minnesota. The purpose of these centers is to develop sequential patterns for teaching reading, composition, and related language skills, to test promising materials, and to develop curriculum recommendations and materials.

Thirty basic or applied research projects related to the improvement of the instruction of English have also been approved.

Funds have been made available for this fiscal year and the next two years for demonstration centers. These centers will disseminate information about particularly successful content, techniques, and programs of instructions. Such centers may be sponsored on the elementary, secondary, or college level.

Individuals or groups may present proposals for studies in any of the above three areas. The proposals should be submitted to the U.S. Office of Education through a college, university, or state department of education. Inquiries should be addressed to: Coordinator of Project English, U.S. Office of Education, Washington 25, D.C.

A RECENT PAMPHLET BY SIDNEY MARGOLIUS, *How to Pay for Your Child's College Education*, reports a number of findings related to the cost of attending college. The average student will spend about $12,000 during a four year period at a prestige college. His counterpart who attends a state college will spend about $6,000 during the same period of time.

College tuition rose 48 percent and the cost of room and board rose 10 to 12 percent between 1958 and 1962.

The rise in college costs is most threatening to students from the low socio-economic groups. Scholarships may not be aiding this group as much as one might suspect. This pamphlet cites a study by Paxton regarding the distribution of National Merit Scholarships. In the one city where Paxton made her study, the students in the upper income constituted 20 percent of the total student population. This upper income group, however, received 67 percent of the National Merit Scholarships.

Readers are invited to send items of research interest to Professor John Schmid, College of Education, University of Arkansas, Fayetteville, Arkansas or to Dr. Willard E. North, Central State College, Warrensburg, Missouri.

Among Those Present

DONALD BARNES, associate professor of education at Ball State Teachers College, holds degrees from Miami, Columbia, and Colorado State. He has contributed to numerous journals in Australia, India, England, and the United States. Recently he has been affiliated part-time with the Midwest Program on Airborne Television Instructions.

LELON R. CAPPS, a Minnesota PhD. is assistant professor of education at the University of California, Santa Barbara. He has taught at all levels of education in Wisconsin, Minnesota, California, and Hawaii.

NORMAN M. CHANSKY, Columbia PhD., is school psychologist with the Ulster County Board of Cooperative Educational Services, New Paltz, N.Y. In association with the National Institute of Mental Health he conducted research on the effects of promotion and retention on slow achievers.

LAWRENCE M. DeRIDDER obtained his bachelor's degree at Northern Michigan College and both his master's and doctor's degrees at the University of Michigan. Between 1950-53 he was a student counselor and assistant professor in psychology at the Undergraduate Division of the University of Illinois and, from 1953 to the present, a member of the College of Education of the University of Tennessee as an associate professor of educational psychology. His articles have appeared in the *Elementary School Journal*, *College and University*, *School and Society*, *Junior College Journal*, *Journal of Higher Education*, and *Journal of Educational Research*.

TRAVIS L. HAWK obtained his bachelor's and master's degree at North Texas State College and his doctorate at The University of Texas. He taught and served as elementary school principal eight years in the public schools of Texas. He has been affiliated with The University of Tennessee College of Education since 1958 as an associate professor of educational psychology.

JOHN C. GLIDEWELL is director of research and development of the St. Louis County Health Department and program director, Social Science Institute, Washington University, St. Louis, where he also serves on faculties of psychiatry, psychology, and sociology. His PhD. is from Chicago. HERBERT DOMKE, a University of Chicago MD and Harvard DPH, is director of the Allegheny County Health Department, Pittsburgh, Pa. MILDRED B. KANTOR is director of vital statistics, St. Louis County Health Department, and assistant director of the community mental health research training program, Social Science Institute, Washington University. Her PhD. is from North Carolina.

ALVIN E. ROBINSON is assistant superintendent for educational services, Oxnard School District, Oxnard, California. His EdD. is from Colorado State. He has been an elementary teacher, elementary principal, and professor of education.

NORMAN E. WALLEN is associate professor of educational psychology at the University of Utah. He received his PhD. from Syracuse. VERNON E. HAUBRICH is associate professor of education at Hunter College in the Bronx. His PhD. is from Wisconsin. IAN E. REID is a doctoral student at Utah.

SHIRLEY ULLMAN WEDEEN, New York University PhD., is assistant professor of education and coordinator of the Basic Skills Center at Brooklyn College. She is the author of *College Reader*, assistant editor of *NART*, and co-editor of *MACE*.

HORTENSE G. WHITE is head of the department of business education at Pearl High School, Nashville, Tenn. She holds a BA from Fisk and a BS from Tennessee A and I, has been a teacher at Nashville since 1944, and is currently an officer in the Nashville City Teachers Assn. EMMA BRAGG is professor of psychology and education at Alabama State College. Her PhD. is from Chicago.